UNIVERSITY
OF MICHIGAN

MAY 5 1955

SOCIAL SCIENCE
LIBRARY

BF
1
J86

Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
AIR FORCE PERSONNEL AND TRAINING RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *AF Personnel and Training Research Center*
W. R. Garner, *The Johns Hopkins University*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *AF Personnel and Training Research Center*
Arthur L. Irion, *Tulane University*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

Effects of Amount, Rate, and Stage of Automatically-Paced Training on Self-Paced Performance: C. O. NYSTROM, R. E. MORIN, AND D. A. GRANT 225
Effect of Rate of Automatically-Paced Training in a Multidimensional Psychomotor Task: N. H. ANDERSON, F. H. KRESSE, AND D. A. GRANT 231
Effect of Target Brightness and Target Speed upon Tracking Proficiency: J. F. VOSS 237
Effect of Frequency of Target Intermittence upon Tracking: W. F. BATTIG, J. F. VOSS, AND W. J. BROGDEN 244
The Frequency Response of Skilled Subjects in a Pursuit Tracking Task: M. NOBLE, P. M. FITTS, AND C. E. WARREN 249
The Influence of Nonreinforcement of a Component of a Complex Stimulus on Resistance to Extinction of the Complex Itself: D. D. WICKENS AND J. D. SNIDE 257
Effect of Varied Reinforcement on Speed of Locomotion: F. A. LOGAN, E. M. BEIER, AND R. A. ELLIS 260
Generalization of Serial Position in Rote Serial Learning: R. W. SCHULZ 267
The Shape of the Discrimination Gradient for Two Intracontinuum Stimulus Separations: G. E. PASSEY AND P. N. HERMAN 273
A Developmental Study of Transposition: H. W. STEVENSON, I. ISCOE, AND C. MCCONNELL 278
The Influence of Past Associations upon Attributive Color Judgments: K. E. BAKER AND I. MACKINTOSH 281
Implicit Verbal Chaining in Paired-Associate Learning: W. A. RUSSELL AND L. H. STORMS 287
The Relationship Between Occipital Alpha Activity and Laterality: A. D. GLANVILLE AND J. J. ANTONITIS 294
Erratum 300

American Psychological Association

Vol. 49 No. 4 April 1955

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1955 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Department of Psychology, Stanford University, Stanford, California, until June 1, 1955. After June 1 address manuscripts to the editor at Air Force Personnel and Training Research Center, Box 1557, Lackland Air Force Base, San Antonio, Texas. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail.

To save printing costs supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 [4, Suppl.], 389–449). In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 25th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Copyright, 1955, by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 49, No. 4 APRIL, 1955

EFFECTS OF AMOUNT, RATE, AND STAGE OF AUTOMATICALLY-PACED TRAINING ON SELF-PACED PERFORMANCE [1]

CHARLES O. NYSTROM, ROBERT E. MORIN, AND DAVID A. GRANT

University of Wisconsin

The present experiment investigates three factors associated with automatically-paced (AP) training on subsequent self-paced (SP) performance of a psychomotor task. These factors are: (*a*) rate of automatic pacing; (*b*) amount of automatic pacing; and (*c*) the locus in the training schedule of the AP practice. In an earlier study (**4**) using the same task, it was observed that, when tested for transfer under a self-paced condition, an AP training group performed poorer than the SP training group on the initial transfer trials but was better on the subsequent blocks of transfer test trials. The AP training group showed significantly greater improvement than the SP training group on the SP transfer test. The same result was reported by Adams (**1**). Since the finding in the earlier study was unexpected, it was desirable to test it in further investigation and to explore the important factors determining this phenomenon.

[1] This research was supported in part by the United States Air Force under Contract Number AF 33(038)-23294 monitored by the Armament Systems Personnel Research Laboratory, Air Force Personnel and Training Research Center, Lowry Air Force Base, Denver, Colorado. Reproduction in whole or in part is permitted for any purpose of the United States Government.

METHOD

Apparatus.—The apparatus was the Multiple Serial Discrimeter (MSD) (**3, 4**) which was designed and constructed at the University of Wisconsin, Laboratory of Experimental Psychology. It consists of five basic units: (*a*) response keyboard, (*b*) stimulus display, (*c*) stimulus programming unit, (*d*) control unit, and (*e*) operations recorder.

The response keyboard consists of a row of eight piano-type lucite keys mounted on the top surface of a gray metal box. The keys are 1 in. apart except for the middle pair which are 3 in. apart. Depression of a key requires 1 oz. of pressure, and operates a microswitch connected with a relay in the scoring circuit and also controls one of the eight response feedback lights on the stimulus display. An adjustable shield on the response keyboard prevents *S* from seeing the keys or his fingers.

The stimulus display consists of two parallel horizontal rows of pilot lights. The upper row are red jewelled stimulus lights and the lower corresponding row are green jewelled response feedback lights. The rows are 1.5 in. apart and the lights within a row are 1 in. apart except that there are 2.5 in. between the middle pair of lights. The left-most stimulus light was associated with the left-most response key, and

successive lights were respectively associated with successive keys.

The stimulus programming unit determines which of the stimulus lights will be lighted on a given trial, thus programming the stimulus patterns. In the present experiment, each stimulus pattern consisted of three lights. The programming is accomplished by Western Union telegraph tape transmitters. Holes punched in the tape operate the stimulus lights and stimulus relays. The tape is advanced to a new pattern whenever a solenoid is activated. Under SP procedures, the solenoid is activated as soon as the previous pattern has been matched. Under AP conditions, the solenoid is activated by an adjustable electronic timing unit.

The control unit consists of the circuitry necessary for stimulus presentation and scoring. Stimulus relays and response relays are wired so that an impulse is passed through the 16 relays if, and only if, *S*'s response corresponds exactly to the stimulus pattern. In self-pacing, this correction impulse advances the transmitter to the next stimulus pattern, but in automatic pacing, when a correct match is made, all stimulus lights are turned off until the next pattern is presented. Two Veeder-Root counters keep separate records of the number of patterns presented and the number of patterns correctly matched.

Response latencies and error indications for each stimulus pattern are recorded by an Esterline-Angus operations recorder.

Subjects.—The *Ss* were 120 male students obtained from a pool of students who had volunteered to serve as part-time paid *Ss*. They were randomly assigned to the 11 experimental procedures under the restriction that each replication be completed before the next one was begun. None of the *Ss* had prior experience with the MSD.

Design.—Each *S* had to match 425 three-light stimulus patterns, identically arranged for all *Ss* in 17 blocks of 25 patterns per block. Patterns were randomly ordered except that each of the 56 possible patterns occurred for the n-th time before any pattern occurred for the $(n + 1)$-th time, and the same pattern was not permitted to occur twice in succession. During the last seven blocks, the transfer test blocks, all groups matched patterns under the SP procedure, which provided a basis for comparing the effects of the earlier training. The experimental treatments consisted of varying the amount, rate, and locus of occurrence of AP practice during the ten training blocks. Two of the ten experimental groups received AP practice during all ten training blocks; they differed in the rate at which patterns were presented to them: i.e., 1.28 sec. per pattern or 1.80 sec. per pattern. The eight other experimental groups resulted from a $2 \times 2 \times 2$ factorial design which involved one or four blocks of AP practice at a fast or slow rate, and administered early (Block 1 or Blocks 1–4) or late (Block 10 or Blocks 7–10). Each of these ten experimental groups consisted of ten *Ss*. The control group, however, contained 20 *Ss*. This group matched all 17 blocks of patterns, training and test, under the SP procedure.

Before the experiment each *S* was carefully instructed as to the procedure of operating the MSD and was given a demonstration. The *Ss* were instructed to "do your best to match each new pattern as quickly as you can. The faster you match the patterns the better your score will be." One-minute rest periods were given after each block of patterns. The *Ss* were always told during the rest period whether the next block of patterns would be automatically- or self-paced. Scores were not given to *Ss* until the end of the experiment.

Results

All scores were reduced to mean seconds to match a pattern. The complete summary of the data of the experiment is presented graphically in Fig. 1. Each subplot in Fig. 1 is a graph of mean seconds to match a pattern as a function of the 17 successive blocks of 25 three-light patterns. The control curve is plotted in each of the five graphs. In addition to the control curves, there are plotted in: (*A*) the two groups receiving AP practice on the first block, (*B*) the two groups receiving AP practice on the tenth block, (*C*) the two groups receiving AP practice on Blocks 1–4, (*D*) the two groups receiving AP practice on Blocks 7–10, and (*E*) the two groups receiving AP practice on all ten training blocks.

Regular learning was obtained for all groups. The regularity of the data permitted a comparison of group performances which indicated that the successive blocks of patterns differed somewhat in difficulty. This is particularly important because Blocks 13 and 14, in the last seven critical test blocks, were unusually easy and

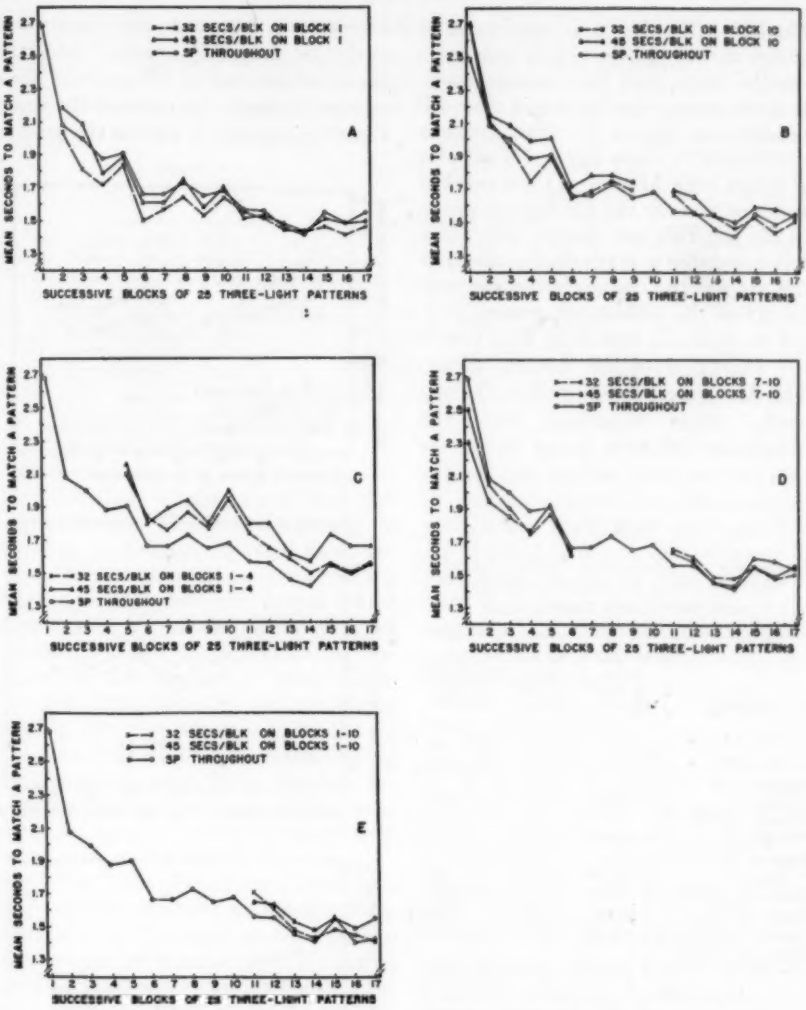

FIG. 1. Self-paced performance curves for all groups. Different subplots show performance for different amounts and stages of AP training.

Block 15 was unusually difficult. In spite of the irregularity due to block difficulty, the same general trend which stood out in the earlier study appeared in this experiment. In general, the SP control group performed better than the AP groups during the initial test blocks, but during the later test blocks the reverse relationship occurred. The major differences among the treatments and their combinations, then, are to be found by comparing the slopes of the group curves during the seven test

blocks. Of the ten experimental group curves in Fig. 1, only one has a smaller slope than the control group; it is the group that received slow AP practice on Block 1. The greatest differences in slope appear in subplot *E* where both AP groups show marked superiority over the SP control group in the last two test blocks. In Table 1 is presented a statistical analysis of the slopes and means over the test blocks of the orthogonal groups plotted in *A*, *B*, *C*, and *D* of Fig. 1 with the exception of the control group. The Alexander trend analysis (**2**) was used. Most important were the differences between group slopes for the fast vs. slow pacing and for the interaction between the factor of early vs. late AP practice and the factor of amount of AP practice (one or four blocks). In general, the more rapid the automatic pacing the greater

TABLE 1

SUMMARIES OF ALEXANDER TREND ANALYSIS OF THE EIGHT ORTHOGONAL GROUPS OVER THE SEVEN TEST BLOCKS

Source	A. Between Group Slopes		B. Between Group Means		C. Group Deviations from Linearity	
	df	*F*	*df*	*F*	*df*	*F*
Stage of AP training	1	0.00	1	0.03	5	0.60
Amount of AP training	1	0.64	1	1.73	5	1.65
Rate of AP Training	1	4.19*	1	0.83	5	2.34*
Stage × Amount	1	4.04*	1	2.58	5	0.40
Stage × Rate	1	0.00	1	0.10	5	0.31
Amount × Rate	1	0.10	1	0.02	5	0.38
Stage × Amount × Rate	1	0.62	1	0.32	5	0.72

Source	*df*	Mean Square	*F*
Between individual slopes	72	.0180	1.78**
Between individual means	72	.4096	40.67***
Over-all slope	1	.7214	71.46***
Over-all deviations from linearity	5	.0221	21.90***
General error estimation	360	.0101	

* Significant at 5% confidence level.
** Significant at 1% confidence level.
*** Significant at .1% confidence level.

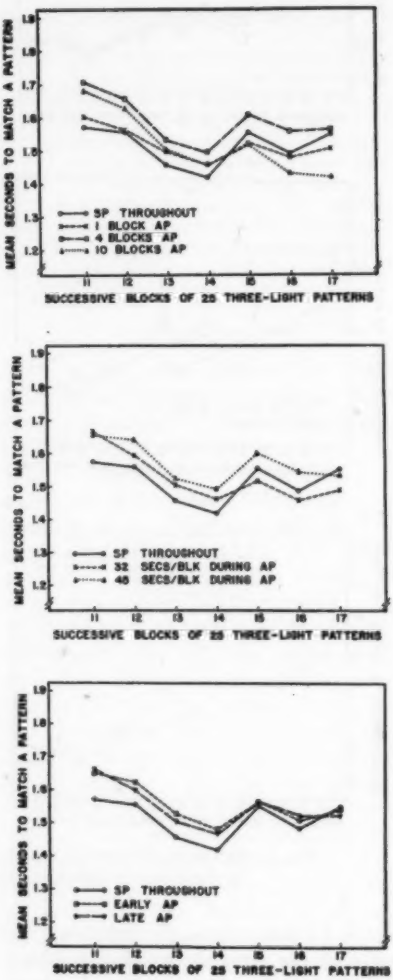

FIG. 2. Each subplot shows the effect of one of the three main variables during the seven test blocks. Top: amount of AP training. Center: rate of AP training. Bottom: stage of AP training.

TABLE 2

SUMMARIES OF ALEXANDER TREND ANALYSES OF ALL 11 GROUPS OVER THE SEVEN TEST BLOCKS

Source	A. 10, 4, 1, and 0 AP Training			B. Fast, Slow AP, and SP Control Group		
	df	*F*	Error Used	*df*	*F*	Error Used
(1) Between group slopes	3	7.48***	(6)	2	6.07**	(6)
(2) Between group means	3	0.74	(7)	2	0.51	(7)
(3) Group deviations from linearity	15	0.66	(8)	10	1.01	(8)
(4) Over-all slope	1	110.07***	(8)	1	111.00***	(8)
(5) Over-all deviations from linearity	5	29.22***	(8)	5	29.47***	(8)
(6) Between individual slopes	116	1.52*	(8)	117	1.66*	(8)
(7) Between individual means	116	37.01***	(8)	117	37.40***	(8)
(8) General error estimate	580	(0.1163)		585	(0.1174)	

* Significant at 5% confidence level.
** Significant at 1% confidence level.
*** Significant at .1% confidence level.

was the slope of the curves over the test blocks. The significant *F* in Column C reflects the fact that the slow AP test trend is more curvilinear than the fast AP test trend. The significant *F*'s in the lower part of Table 1 show the existence of individual differences and progressive but nonlinear decreases in time scores during the test blocks. The *F*'s in Column A involved *Between individual slopes* as the error term; *Between individual means* was used for Column B; and Column C involved the general error term.

In Fig. 2 are plotted for comparison the transfer test trends for the basic variables of the experiment. In the top graph are compared the transfer test curves for the various amounts of AP training (0, 1, 4, or 10 blocks) during the experiment. The curves show progressively greater slopes for increasing amounts of AP practice. In the center graph the two rates of automatic pacing are compared with the SP control curve. Here is a clear indication of greater improvement in the fast AP training (32 sec. per block) as compared with the slow AP training (45 sec. per block) and the SP control. In the bottom graph the comparison is between early and late AP training. Here the presence of automatic pacing in the training is shown to increase the slope of the curves, but there is no apparent difference between early and late automatic pacing.

The trend analyses of the data of the top and center subplots of Fig. 2 are presented in Table 2. In Column A of Table 2 a comparison is made between the various amounts of AP training—no automatic pacing in the control group, 1 block, 4 blocks, and 10 blocks in the experimental groups. All 11 groups were used in this analysis. There is a significant difference between group slopes indicating that the more AP training given the more rapid is the improvement during the test blocks as compared with the SP control group. Other significant differences indicate that there is over-all learning with both linear and nonlinear components, and that there are individual differences within the groups. The *F*'s in Column B are for the comparison between the fast, slow, and control groups. All 11 groups were analyzed. Again, there

is a significant difference between group slopes with an F of 6.07. The indication is that the more rapid automatic pacing is associated with greater improvement during the test blocks as compared with the control group performance. And again, there is indication of over-all slope and over-all deviation from linearity.

Discussion

These findings confirm and extend the observations reported in the earlier study in this series and are in line with the results reported by Adams (1), although Adams did not separately test the significance of the linear and nonlinear components of his trends. The indications are clear that automatic pacing in training can improve later SP performance. This improvement is observed in terms of nine of the ten experimental groups having greater slopes than the control group during the test blocks. In nearly all cases this resulted in the superiority of the experimental groups over the control group in the last few test blocks.

No significant differences between group means were observed. Despite this, the findings are important for two reasons. First, it is surprising that AP groups even do as well during the test blocks as the SP group, for the AP *S*s match fewer patterns correctly and in this sense received less practice than the SP *S*s during the training blocks. Secondly, if the test block slopes were to remain relatively constant, then later SP performance would be benefited more by AP training than by SP training itself.

Summary

Eleven groups of *S*s were given ten blocks of training on the MSD psychomotor task, then transferred to a self-paced procedure for the remaining seven test blocks. Their training permitted study of the effects on SP performance of three aspects of automatic pacing: (*a*) amount (0, 1, 4, or 10 blocks), (*b*) rate (1.28 vs. 1.80 sec. per pattern), and (*c*) stage of AP practice (early vs. late). Within the limits studied, the principal (statistically significant) findings were as follows:

1. The greater the amount of AP practice given during training, the greater the rate of improvement during the subsequent SP test blocks.
2. The faster the AP practice, the greater the rate of improvement during the test blocks.
3. There was no significant effect due to early vs. late AP practice on the transfer test performance, except that there was a slope interaction between stage and amount of AP practice.

The present results confirm earlier findings (**1, 4**) and reveal more of the extent to which the variables of amount and rate of AP training are beneficial to later SP performance.

References

1. Adams, J. A. The effect of pacing on the learning of a psychomotor response. *J. exp. Psychol.*, 1954, **47**, 101–105.
2. Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.
3. Morin, R. E., & Grant, D. A. Learning and performance on a key-pressing task as a function of the degree of spatial stimulus-response correspondence. *J. exp. Psychol.*, 1955, **49**, 39–47.
4. Nystrom, C. O., Morin, R. E., & Grant, D. A. Transfer effects between automatically-paced and self-paced training schedules in a perceptual-motor task. *USAF, Hum. Resour. Res. Cent., Res. Bull.*, 1953, No. 53–66.

(Received July 1, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

EFFECT OF RATE OF AUTOMATICALLY-PACED TRAINING IN A MULTIDIMENSIONAL PSYCHOMOTOR TASK [1]

NORMAN H. ANDERSON, FREDERICK H. KRESSE, AND DAVID A. GRANT

University of Wisconsin

The present experiment investigates the effect of rate of automatically-paced (AP) training on subsequent self-paced (SP) performance in a psychomotor task. Nine conditions of AP training were employed in addition to the SP control group. Six of the AP groups were trained at constant pacing rates. For the three remaining experimental groups, the pacing rate was increased in successive training trials. The variable rate for one of these groups followed the average SP performance curve on a previous study with the present apparatus (**3**), while the remaining two pacing speeds lay 10% above and 10% below this curve.

Previous studies have shown that there is excellent transfer from AP training to the SP test. Adams (**1**), using a modified Mashburn task, found no significant differences in SP performance times between control and experimental groups equated on total training time. In experiments involving a finger keyboard response (**4, 5**), it was found that groups trained on AP showed greater improvement on the SP test than did the SP control group. Rate and amount of AP training were found to be important variables. The present experiment was carried out using a different psychomotor task in order to obtain further information on the generality of the earlier findings.

[1] This research was supported in part by the United States Air Force under Contract No. AF33(038)-23294, monitored by the Armament Systems Personnel Research Laboratory, Air Force Personnel and Training Research Center, Lowry Air Force Base, Denver, Colorado. Reproduction in whole or in part is permitted for any purpose of the United States Government.

METHOD

Apparatus.—The apparatus used in this experiment was designed to provide an approximate simulation of radar procedures, although no particular radar was simulated. A detailed description of its operation has been given in an earlier paper (**3**) so that only a brief résumé need be given here.

The *S* sat facing a 6-in. oscilloscope screen divided symmetrically into four 2-in. square quadrants. Stimulus patterns appeared in the centers of these quadrants, and *S*'s task was to center these patterns on the scope with one-handed corrective movements of a control stick mounted on a pedestal placed between *S*'s legs.

The patterns were presented by a programming unit consisting of a Western Union tape transmitter and associated relays and electronic circuits. Under SP presentation, a new pattern appeared as soon as *S* had correctly responded to the current pattern. The AP presentation was governed by an electronic timing circuit which presented new stimuli after a certain time interval regardless of *S*'s actions. In this case, whenever *S*'s response was correct, the stimulus pattern locked in the center of the scope face until the end of the allotted time. Errors for elevation, azimuth, and range, as well as hits, were scored by four counters.

The stimulus patterns were long (½ in.) or short (¼ in.) horizontal lines appearing in the center of one of the four quadrants on the oscilloscope screen. The *S*'s task was to produce a medium sized (⅜ in.) line in the center of the screen by making three corrective motions with the stick. For all patterns the effects of the stick motions on the patterns were as follows:

Azimuth: Stick right moved stimulus left,
stick left moved stimulus right;
Elevation: Stick forward moved stimulus up,
stick backward moved stimulus down;
Range: Thumb control forward lengthened stimulus,
thumb control backward shortened stimulus.

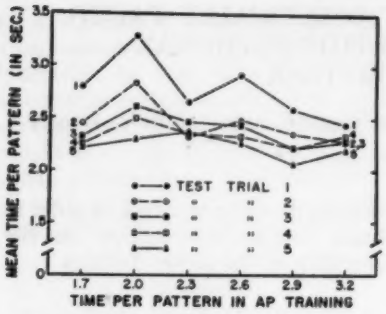

FIG. 1. Variation of mean response time per pattern in the transfer test as a function of time per pattern in AP training for the six constant-rate AP groups. The parameter is test trial number.

In addition to this visual response feedback, a disagreeable buzzer sounded whenever *S* made an error.

Experimental design.—Ten groups were run, each containing ten *S*s. Each group received a total of 14 trials of 40 stimulus patterns each. The first 9 of these trials were used for training and the last 5 were used to test the effects of training. All groups received the initial training trial and the 5 test trials under SP; they were differentiated by their treatment on the last 8 trials of the training series. Six groups received these 8 trials at one of six constant rates of stimulus presentation. Three additional groups were also run under AP training but successive trials were presented at faster rates. A control group was given SP training throughout.

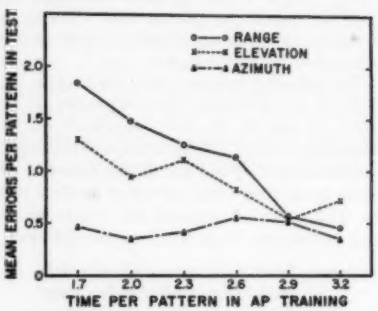

FIG. 2. Variation of mean errors per pattern in the transfer test as a function of time per pattern in AP training for the six constant-rate AP groups.

For the six constant-rate AP groups, each stimulus was presented for a duration of either 1.7, 2.0, 2.3, 2.6, 2.9, or 3.2 sec. For one shifting-rate AP group, the pacing speed followed the average SP performance curve in the earlier experiment (3), while the pacing speed for the remaining two groups lay 10% above and 10% below this average curve. Comparison showed that the control group performance in the present experiment coincided quite well, after the first three training trials, with the performance predicted by the earlier experiment, though at the outset their performance was considerably inferior.

Procedure.—After being seated at the apparatus, *S* was read the instructions, identical for all groups, in the course of which he corrected two stimulus patterns. After answering questions, *E* then presented the first trial of 40 stimuli using the SP procedure. In the 1½-min. interval which followed completion of this trial, the experimental groups were informed that the stimuli would now be presented using the AP procedure, with appropriate wording for the constant-rate and for the shifting-rate groups. To reduce possible frustration, these groups were told specifically not to worry if the stimuli seemed to come on too fast, but to do as well as possible. There followed eight trials of 40 stimuli each, with a 1-min. rest interval between successive trials. After the training series, all *S*s were given a 5-min. break before the beginning of the test series. In the last ½ min. of this rest, *E* read a brief set of instructions informing *S* that the test series was about to begin, that it would consist of five trials of 40 stimuli each, and that the SP presentation would be used. Two *E*s conducted the experiment, each running half of the *S*s in all groups.

Subjects.—The *S*s were 100 male undergraduate students who had signed up as paid *S*s for psychological experiments. None had worked on previous studies with this apparatus.

RESULTS

The results to be presented in the present paper are restricted to the transfer test at the end of training. The scores used are mean time per stimulus averaged over the 40 stimuli of each test trial for each *S*, and mean number of errors per stimulus pattern averaged over the five trials of the test series.

Effect of pacing rate in AP training on subsequent SP test.—Mean response

time per stimulus pattern in the SP transfer test is plotted against pacing rate in AP training in Fig. 1. The parameter is test trial number. The graph gives some indication of a decrease in time per pattern with increase in training time. However, analysis with orthogonal polynomials showed that this trend was not significant.

In Fig. 2, number of errors per stimulus pattern averaged over the 200 patterns of the five test trials are graphed separately for range, elevation, and azimuth, with pacing rate in AP training as the abscissa. Range errors decrease sharply as AP training time increases. Elevation errors also show some decrease but azimuth errors are essentially equal for the various AP groups. These statements are supported by the orthogonal polynomial analysis given in Table 1.

Trend during transfer test.—Figure 3*a* compares average performance of the six constant-rate AP groups and the SP control group for each of the five test trials. Inclusion of the control group performance on training Trials 2 through 6 indicates the extent of the benefit conferred by AP training.

There is a clear tendency for the faster AP groups to show steeper

TABLE 1

SUMMARIES OF ANALYSIS OF RANGE, ELEVATION, AND AZIMUTH ERRORS ON SP TEST AS A FUNCTION OF AP TRAINING RATE

Source of Variation	*df*	A. Range Errors	B. Elevation Errors	C. Azimuth Errors
Between rates	5	3.07*	0.80	0.49
Linear component	1	14.86***	2.98	0.00
Residual	4	0.13	0.25	0.61
Within rates				
Error estimate	54	*362.50*	*369.43*	*59.88*

* Significant at 5% confidence level.
*** Significant at .1% confidence level.

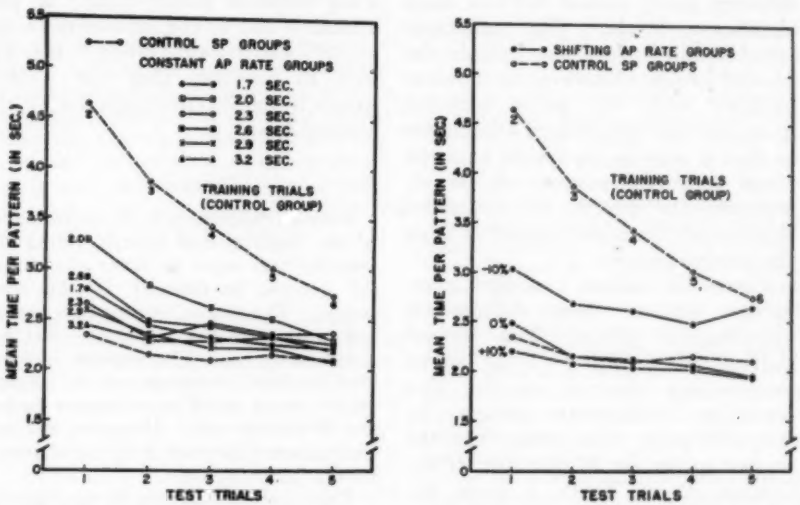

FIG. 3. Variation of mean response time per pattern in the transfer test as a function of test trial number: (*a*) control SP group compared with the six constant-rate AP groups; (*b*) control SP group compared with the three shifting-rate AP groups. In both cases the upper curve gives control group performance on training Trials 2 through 6.

TABLE 2

ANALYSIS OF TRENDS OF TIME SCORES DURING SP TESTS FOLLOWING VARIED RATES OF AP TRAINING

Source of Variation	A. Constant AP Rates		B. Shifting AP Rates	
	df	*F*	*df*	*F*
1. Between group means	5	0.72	2	4.64*
2. Between group slopes	5	3.54**	2	0.34
3. Between individual means	54	41.20***	27	21.36***
Error estimate for (1)		*2610.89*		*2025.33*
4. Between individual slopes	54	3.34***	27	4.43***
Error estimate for (2)		*211.45*		*420.39*
5. Group deviations from estimation	15	0.68	6	1.25
6. Over-all slope	1	215.88***	1	41.08***
7. Over-all deviations from linearity	3	13.76***	3	4.38**
8. Individual deviations from estimation *Error variance estimate*	162	*63.37*	81	*94.83*

* Significant at 5% level.
** Significant at 1% level.
*** Significant at .1% level.

slopes in this graph, and this was borne out by the Alexander trend analysis (**2**), given in Column A of Table 2, which shows significant (.01) differences between group slopes. At the same time, differences between group means did not reach significant levels. The tabulated trend analysis does not include the control group. However, a separate analysis with this group included gives identical significance statements so that it may be concluded that the mean test performances of the six constant-rate groups do not differ significantly from one another or from the control group.

Figure 3*b* makes a similar comparison for the three shifting-rate experimental groups. The slowest shifting-rate group (+10%) shows performance superior to the two remaining shifting-rate groups. In fact, this group does better than the control group on all five test trials. Column B of Table 2 gives the Alexander trend analysis of these three groups. Here it is seen that group slopes are not significantly different but that the difference between group means is significant (.05). When the control group data were added, the between group means *F* was reduced somewhat, but the same significance levels were obtained for all group comparisons. This difference between group means is presumably due to the inferiority of the − 10% shifting-rate group. It is not safe to conclude that the + 10% group is significantly superior to the control group.

DISCUSSION

Over-all differences in the performance of the experimental groups during the transfer test seem to favor the slower AP groups, particularly in the error scores. The slower, shifting-rate groups are superior to the SP control but not significantly so. The results indicate that the faster constant-rate AP groups show a more rapid improvement during the SP transfer test. However, this just compensated for their poor initial transfer test performance.

These conclusions are in contrast to the results of the finger keyboard study (**4, 5**), where AP training eventuated in performance superior to that of the SP group. This difference could arise in

two ways. First of all, intrinsic differences in the two tasks could lead to these slight discrepancies. Certainly, not enough is known to state in what way, and to what extent, two psychomotor tasks are equivalent. Secondly, it is possible that AP training confers a real benefit, but the length of the transfer test was insufficient to reveal it.

It is also of interest to compare the present data with Adams' modified Mashburn task (1). In both cases, no significant differences in SP performance times were found between the groups given AP training and the SP control group despite marked differences in numbers of patterns successfully responded to in the training series. In Adams' study, however, the various groups were equated on training time which was not the case in the present research. These differences in training time are reflected in Fig. 4, which plots, at each trial of the test series, total prior task time for each group relative to the control group. Thus, for each group, the points at Trial 1 give the total time spent on the nine training trials divided by 1271 sec., which is the average total training time for the control *S*s. Except for the slowest pacing rate in each type of AP training, all curves lie below the control group line, showing that, at each trial, these groups had spent less time at the task than had the control group. Although this graph must be interpreted with caution, it indicates that, in terms of relative training time, AP training is more efficient than SP training.

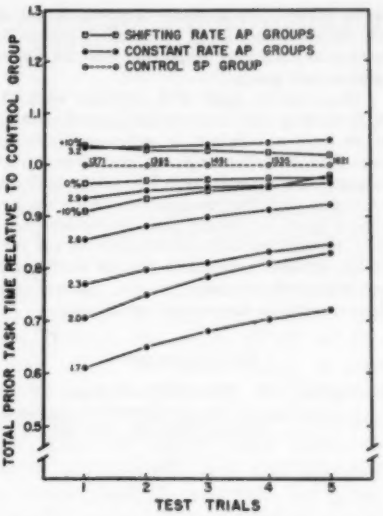
SHIFTING RATE AP GROUPS
CONSTANT RATE AP GROUPS
CONTROL SP GROUP
TOTAL PRIOR TASK TIME RELATIVE TO CONTROL GROUP
TEST TRIALS

FIG. 4. Variation of total prior task time relative to the control group as a function of test trial number. Numbers by control group points give mean total task time in seconds for the control group prior to the given trial. Numbers by the AP group curves indicate pacing rate in training.

SUMMARY

The results of an experiment testing the effect of rate of automatically-paced training in a multidimensional psychomotor task are presented. Nine conditions of automatic pacing were employed in addition to the self-paced control group. Six of the experimental groups were trained with constant pacing rates while the pacing rate for the remaining three experimental groups was increased in successive training trials.

The experimental *S*s were given the initial training trial under self-pacing followed by eight training trials under automatic pacing. All groups were given a transfer test consisting of five trials under the self-paced procedure. Each trial contained 40 stimulus patterns. Time scores and errors in range, azimuth, and elevation are reported for the five trials of the test series.

The following transfer test results were obtained for the constant-rate automatically-paced groups: (*a*) Range errors decreased sharply as time per pattern in automatically-paced training increased. Elevation errors also showed a decrease, while azimuth errors were essentially equal for all groups. (*b*) A tendency toward decreased response time for the slower pacing rates was found to be not significant. (*c*) An Alexander trend analysis showed no difference in mean performance times for these groups either alone or together with the control group. Differences between group slopes were significant indicating that the groups trained under the faster pacing rates were improving more rapidly in the transfer test.

The following transfer test results were obtained for the shifting-rate automatically-paced groups: (*a*) No significant differences between group slopes were found by the trend analysis, either for these three groups alone, or when combined with the control group. (*b*) Differ-

ences between group means were significant at the .05 confidence level. This effect seemed to be due to the poorer performance of the fastest shifting-rate group.

These results agree with previous work (1, 4, 5) showing that automatically-paced training is at least as efficient as self-paced training, even though the former groups give fewer correct responses during training. The results indicate, moreover, that this remains true even when less time is spent in training than for the self-paced *Ss*. Superior terminal performance on the part of the automatically-paced groups, as reported in two previous researches (4, 5), was not demonstrated in the present experiment.

References

1. Adams, J. A. The effect of pacing on the learning of a psychomotor response. *J. exp. Psychol.*, 1954, **47**, 101–105.
2. Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.
3. Kresse, F. H., Peterson, R. M., & Grant, D. A. Multiple response transfer as a function of supplementary training with verbal schematic aids. *J. exp. Psychol.*, 1954, **48**, 381–390.
4. Nystrom, C. O., Morin, R. E., & Grant, D. A. Transfer effects between automatically-paced and self-paced training schedules in a perceptual-motor task. *USAF, Hum. Resour. Res. Cent., Res. Bull.*, 1953, No. 53–66.
5. Nystrom, C. O., Morin, R. E., & Grant, D. A. Effects of amount, rate, and stage of automatically-paced training on self-paced performance. *J. exp. Psychol.*, 1955, **49**, 225–230.

(Received July 1, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

EFFECT OF TARGET BRIGHTNESS AND TARGET SPEED UPON TRACKING PROFICIENCY[1]

JAMES F. VOSS

University of Wisconsin

A variable which may have a significant effect upon visual tracking is the contrast in brightness of the target being tracked and its background. Battig, Nagel, Small, Gregg, and Brogden (**2**), in discussing their study of the effect of frequency of intermittent presentation of the target upon tracking, indicate that brightness may be important in determining proficiency of tracking. Contrast is the primary variable of the present experiment and was controlled by maintaining a dark background and varying the brightness of the target.

When the brightness of the target is decreased to a point where it cannot be discriminated against the background, a limiting condition of zero performance is established. The relative level of tracking performance for levels of target brightness higher than the limiting condition is unknown. One hypothesis is that, as target brightness increases, tracking proficiency will increase in a linear manner. A second hypothesis asserts that, as target brightness increases, tracking proficiency will increase in a curvilinear manner. A third hypothesis is that tracking proficiency is unaffected by target brightness except at a point where the target border and background cannot be readily discriminated. Acquisition of tracking was studied at each of six brightness levels within the range of 697 mL to .0028 mL. The latter value is just above the threshold of visibility of the target. Transfer trials employing the brightest target were given in order to determine how performance at this level is affected by prior practice on the lower levels of target brightness.

The effect of target speed was also investigated to determine whether speed is a significant variable in tracking proficiency and whether it has a differential effect upon proficiency at different levels of target brightness. If tracking is considered a discrete set of movements, i.e., *S* overtakes the target, passes it, returns, etc., then it may be hypothesized that this pattern of movement would allow *S* to be on target more consistently for a slow speed than for a faster speed. Moreover, because of the possible differences in visibility of the target at different levels of brightness, there may be an interaction between target speed and target brightness.

METHOD

Apparatus.—The apparatus used was the Pedestal Sight Manipulation Test, hereafter referred to as the PSMT. The two PSMT's used in this study are the Mast modification (**8**) of the Type B apparatus described by Melton (**6**). They were located in the same room but separated by a partition.

A brief description of the apparatus is given

[1] This article is based on a thesis presented in partial fulfillment of the requirements for the M.S. degree, Department of Psychology, University of Wisconsin, 1954. The author wishes to thank Dr. W. J. Brogden, who directed the research. This research was supported in part by the United States Air Force under Contract Number AF 33(038)-23294 monitored by the Armament Systems Personnel Research Laboratory, Air Force Personnel and Training Research Center, Lowry Air Force Base, Denver, Colorado. Reproduction in whole or in part is permitted for any purpose of the United States Government.

as follows. The image of the front view of an airplane is projected upon a screen in a darkened room. The image moves across the screen, changing in elevation, wingspan, and in the inclination of the axis of the wingspan. An interpretation of the latter two movements is that the plane is changing in range and bank, respectively. The *S* manipulates a sight by grasping one side of it with each hand, and moving it from side to side, and up and down. There is a ring of dots and a center dot projected from the sight, and *S*'s task is to keep the center dot on the nose of the plane while he keeps the diameter of the ring of dots equal to the wingspan. The size of the reticle is changed by rotation of the range knob which *S* grasps with the fingers of his right hand. The screen is 60 in. high and 93 in. wide.

Subjects.—There were 96 male *S*s, all students at the University of Wisconsin. Each was paid 85 cents per practice session.

Procedure.—In preliminary investigation, the brightness at which "something" could be seen was ascertained. This threshold was obtained by inserting filters of various densities in front of the projector. Six brightness levels were then chosen for the experiment, ranging from the filter level of 5.4 log units, which was approximately the threshold of visibility, to the brightness level where no filter was used. Four intermediate values were also employed. The six target brightness values are 697 mL, 6.97 mL, .0697 mL, .00697 mL, .0044 mL, .0028 mL. Each of six equal-sized groups practiced at one of the six brightness levels. The groups are numbered 1 through 6, in order of decreasing brightness. There are more values at the lower brightness levels since one of the hypotheses tested is that tracking proficiency is not affected until a point where there is practically no contrast. Because the brightness level for Groups 4, 5, and 6 was low, *S*s were required to dark adapt for approximately 10 min. before entering the darkened experimental room.

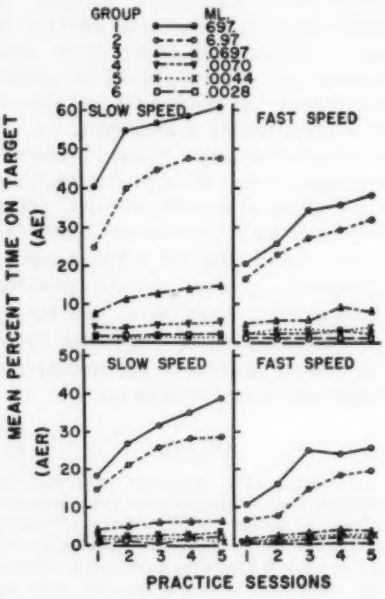

FIG. 1. Tracking performance as a function of practice and target speed.

Each *S* practiced for five consecutive days at the brightness level to which he was assigned. A trial is defined as a series of eight flights. Each day *S* had eight trials with a 1-min. rest after Trials 1, 2, and 3, a 5-min. rest after Trial 4, and a 1-min. rest after Trials 5, 6, and 7. One-half of each brightness group was assigned to each PSMT. Speed was orthogonal to both apparatus and brightness. The time for the simulated airplane to cross the screen at the slow speed was approximately 6.7 sec., and for the fast speed approximately 4.4 sec. The *S*s were assigned randomly with the limitation that the two *S*s finishing together on the fifth day have the same speed. Following a 2-min. rest after Trial 8, *S* was given two additional transfer trials on the fifth day at the highest brightness level. There was a 1-min. rest between the two trials. The *E* recorded time-on-target scores for each trial for azimuth (A), elevation (E), range (R), azimuth and elevation (AE), and azimuth, elevation, and range (AER). The scoring area for range was 16 mils, or approximately $1\frac{1}{16}$ in. The scoring areas for azimuth and elevation were 20 mils, or approximately $1\frac{1}{4}$ in. Calibration was done every third day in order to maintain the above scoring areas.

The instructions were given to *S* only on the first day. They read as follows: "This is a test of your ability to use a B-29 Pedestal Gunsight. Notice that the sight reticle (*E* demonstrates) has been projected upon the screen. Observe the center dot. The ring of dots gets larger as this outer knob is turned forward (*E* demonstrates) and smaller as it is turned backward (*E* demonstrates). Your job will be to keep the center dot on the nose of an airplane while you keep the ring of dots the same size as the plane's wingspan. By grasping the left knob (*E* demonstrates) the sight can be moved from side to side (*E* demonstrates), and up and down (*E* demonstrates). Your score will be the total amount of time you keep the center dot on the nose of the plane while you keep the ring of dots

the same size as the plane's wingspan. You will have several series of attacks with the plane starting alternately from the right and left of the screen. Kneel on the bench. You will be able to score only when the plane is moving. When the red light on your left goes on, that signifies the end of a series of attacks. You are then to let go of the sight and relax. Shortly after the red light goes off, the plane will again go across the screen. Please do not talk to the other person in the room. You will have a minute rest after the first three series of eight flights, then a 5-min. rest after the fourth series, and a minute rest after each of the next series."

Either following their first day of practice, or on the second day, *S*s were told the meaning of the clock scores. They were given knowledge of results in order to promote efficient learning.

Results

Figure 1 presents the acquisition curves for the different target brightness values for the AE and AER scores for each speed. The scores are given as percentages of maximum possible time on target. Since the results for the A, E, and R scores are quite similar to those for the AE and AER scores, only the latter two scores were analyzed. It should be noted, however, that comparison of the A and the E scores shows that a higher level of proficiency is attained in elevation than in azimuth tracking. When the A, E, and AE scores are inspected, it is observed that *S*s were often on target in either azimuth or elevation, but not both. This latter effect exists especially for the groups performing at the lower levels of target brightness.

Separate analyses were performed on the AE and AER scores for both speeds. The assumption of homogeneity of variance was tested for each set of scores by Bartlett's test (4). This assumption was rejected for all of the scores, using the 5% level of confidence. A logarithmic transformation was performed, and the assumption of homogeneity of variance was again tested. It is rejected for both of the AER scores and the AE slow speed score, using the 5% level of confidence. For each of the four sets of scores, the logarithmic transformation greatly reduces the χ^2 value. Norton (5) has found that the effect of nonhomogeneity is not of great importance unless the nonhomogeneity

TABLE 1

Summaries of the Modified Alexander Trend Tests for all Scores

Source of Variation	*df*	V_2**	AE		AER		AE		AER	
			Slow Speed Score *MS*	*F*	Slow Speed Score *MS*	*F*	Fast Speed Score *MS*	*F*	Fast Speed Score *MS*	*F*
1. IDFE	108		2,848		4,426		2,502		4,975	
2. GDFE										
a. Apparatus	3	1	1,047	0.36	1,638	0.37	740	0.29	6,634	1.33
b. Brightness	15	1	7,920	2.78*	12,038	2.72*	5,293	2.12	2,997	0.60
c. A × B	15	1	5,071	1.78	3,212	0.73	2,147	0.86	3,489	0.70
3. BGM										
a. Apparatus	1	7	796,608	15.19*	4,284	0.06	35,042	1.05	74,167	1.60
b. Brightness	5	7	13,662,920	260.49*	11,049,792	153.35*	7,630,763	229.36*	5,710,111	122.90*
c. A × B	5	7	93,473	1.78	64,810	0.90	51,417	1.55	101,985	2.19
4. BGS										
a. Apparatus	1	8	1,864	0.34	1,191	0.08	6,042	1.26	7,715	1.55
b. Brightness	5	8	15,990	2.94	88,308	5.86*	34,583	7.22*	113,872	22.89*
c. A × B	5	8†	7,655	1.41	3,834	0.25	8,409	1.76	9,898	1.99
5. ODFL	3	1	27,362	9.61*	36,941	8.35*	18,559	7.42*	18,444	37.07*
6. OS	1	1	794,139	278.84*	834,834	188.62*	57,602,442	23,022.56*	683,502	137.39*
7. BIM	36	1	52,450	18.41*	72,055	16.28*	33,270	13.30*	46,463	9.77*
8. BIS	36	1	5,433	1.91*	15,069	3.40*	4,791	1.91*	7,200	1.45

* Significant at .01 level.
** Denominator in *F* ratio.
† 1.0 for the AER fast speed score.
Note.—IDFE = individual deviations from estimation; GDFE = group deviations from estimation; BGM = between group means; BGS = between group slopes; ODFL = over-all deviation from linearity; OS = over-all slope; BIM = between individual means; BIS = between individual slopes.

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE FOR THE AE AND AER SCORES WITH CONVERTED SLOW SPEED SCORE

Source of Variation	*df*	AE		AER	
		Mean Square	*F*	Mean Square	*F*
Brightness	5	18,512,212	502.78*	13,715,658	259.80*
Speed	1	1,733,404	47.08*	1,223,110	23.19*
Apparatus	1	483,984	13.14*	40,737	0.77
Brightness × Speed	5	253,618	6.89*	192,911	3.65*
Brightness × Apparatus	5	58,678	1.59	142,974	2.86
Apparatus × Speed	1	185,642	5.04	33,601	0.64
A × B × S	5	47,671	1.29	8,029	0.15
Between *Ss* within groups	72	36,820		52,793	
Trials	4	348,777	124.12*	400,626	9.76*
Trials × Brightness	20	17,104	6.09*	51,778	1.26
Trials × Apparatus	4	2,614	0.93	6,288	0.15
Trials × Speed	4	602	0.02	3,758	0.09
T × B × A	20	3,007	1.07	6,313	0.15
T × B × S	20	4,731	1.68	3,789	0.09
T × A × S	4	596	0.02	4,123	0.10
T × A × S × B	20	4,568	1.62	4,035	0.10
Ss pooled within trials	288	2,810		41,032	

exists in great magnitude. In order to compensate for the possible effect of nonhomogeneity of variance, the 1% level of significance is adopted for all subsequent tests. This confidence level is regarded as equivalent to the 5% level normally employed.

An extension of the Alexander Trend Test which provides for consideration of two variables (1) was performed on each of the four transformed sets of scores. Table 1 summarizes the results of the analyses performed on the AE slow speed, AER slow speed, AE fast speed, and AER fast speed scores.

The BIM mean square is tested against the IDFE mean square. This test was done in order to determine the appropriate error term to test the BGM mean square. In all four analyses, the BIM mean square is significant. The BGM mean square for brightness was found significant in all four analyses when tested with the BIM mean square. The BGM mean square for apparatus was significant for the AE slow speed score. The BIS mean square is tested against the IDFE mean square in order to determine the appropriate error term to test the BGS mean squares. For all scores except the AE slow speed score, the BGS mean square for brightness was significant. The GDFE mean square for brightness is significant for the AE and AER slow speed scores. The OS mean square and the ODFL mean square are significant for all four analyses.

The Duncan range test (3) was applied in order to determine which brightness group means differed significantly from each other. The tests indicate that no significant difference exists between the means of Groups 1 and 2. It is also found that for all scores except the AER fast speed score, Group 3 differs significantly from all other groups. In all scores except the AER fast speed score, Groups 4, 5, and 6 do not differ significantly from each other.

Because the fast speed airplane was

on the screen for .66 of the time that the slow speed airplane was on the screen, the raw scores for the slow speed were multiplied by .66 in order to make the measures of the two speeds comparable with respect to the maximum possible time on target. A logarithmic transformation of both the AE and AER scores was performed, and an analysis of variance was completed on each of these sets of scores (4). These analyses are primarily concerned with whether or not there is a significant difference between speeds and are summarized in Table 2. The results indicate there is a significant difference between speeds for both the AE and AER scores. The brightness groups by speed interaction is also significant for both scores. There is also a significant difference between trials and between brightness groups for each score. The trial by brightness group interaction is significant for the AE scores, as is the difference between apparatus groups.

Figure 2 depicts tracking proficiency as a function of brightness level. The group mean AE and AER scores for the fifth practice session are plotted for each speed. Analysis of these curves was performed in the manner outlined by Snedecor (7). This procedure provides a method of determining whether the function deviates significantly from linearity. The results show that for all four sets of scores the deviation from linearity mean square is significant. Because a significant practice session by brightness group interaction was previously found significant for the AE scores, the test was carried out on the first, second, third, and fourth practice sessions for the AE scores. The deviations from linearity mean square is significant in each of these analyses.

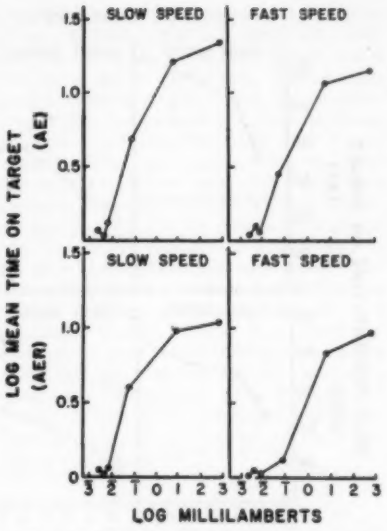

FIG. 2. Terminal tracking performance as a function of target brightness and target speed.

The shape of the function is approximately the same for all five days.

The results of the two transfer trials are shown in Fig. 3 where percentage mean time on target is shown as a function of brightness level for all four sets of scores. The dotted line represents the reference for the measure of transfer and is the mean of the first two trials of Group 1 on the first day. The score for this group and the mean transfer scores of Groups 2, 3, 4, 5, and 6 were analyzed by a simple between-within analysis of variance. The between group mean square is significant for all four sets of scores. A Duncan range test was applied to determine which groups show a significant amount of transfer. Group 2 shows significant positive transfer for all scores. Group 3 shows significant positive transfer for all scores except the AE fast speed score. Groups 4, 5, and 6 do not show a significant amount of positive transfer except for

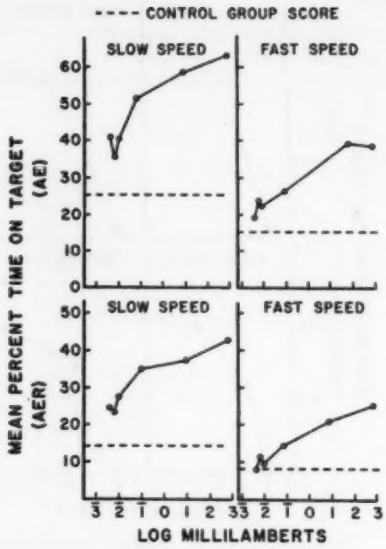

FIG. 3. Transfer of tracking performance as a function of training target brightness and target speed.

the AE slow speed score for Groups 4 and 6.

Another measure of transfer may be obtained by comparison of the performance of the group which practiced at the brightest level with the performance of the groups which practiced at the other brightness levels. This comparison is made in Fig. 3 for each of the four scores by comparing the point for Group 1 (approximately log 3 mL, on the graph) with each of the other points, Groups 2, 3, 4, 5, and 6 (approximately log 1 mL, log $\bar{1}$ mL log $\bar{2}$ mL, log $\bar{3}.8$ mL, and log $\bar{3}.6$, respectively). A Duncan range test for each of the four scores was performed to determine which groups differed significantly from Group 1. The results of the analyses show that for the AER slow speed score Group 1 does not differ significantly on the transfer trials from Groups 2 and 3. Group 1 does differ significantly from Groups 4, 5, and 6. For the fast speed scores and the AE slow speed scores, Group 1 does not differ significantly from Group 2 on the transfer trial, but does differ significantly from Groups 3, 4, 5, and 6.

DISCUSSION

The results of the Alexander Trend Tests and the subsequent Duncan range tests will be discussed first. The significant difference between brightness groups slopes (see Fig. 1) indicates that the rate of learning is faster for the groups practicing at the higher brightness levels. The over-all deviation from linearity was found significant, which indicates that the over-all acquisition curve is not linear. The significance of the groups deviation from estimation term in two of the four scores indicates that the brightness groups differ significantly about their own regression line. The significance of the between brightness groups mean square and the Duncan range tests demonstrates that the groups fall into three brackets, namely, Groups 1 and 2, Group 3, and Groups 4, 5, and 6. These results indicate that tracking performance decreases as target brightness decreases, but only for large decrements in brightness.

The analyses of the data presented in Fig. 2 indicate that there is a significant deviation from linearity when performance is considered as a function of brightness level. From these analyses and from inspection of Fig. 2, it is observed that tracking proficiency increases in a curvilinear manner as target brightness increases. This statement applies to the brightness values used in this experiment. For all stages of practice, it is noted that the hypothesis which states that tracking proficiency will increase in a curvilinear manner as target brightness increases is substantiated by this experiment. For levels of brightness higher than those used in this experiment, tracking proficiency would be expected to increase relatively little. This is ascertained by inspection of Fig. 2 where it is observed that performance

appears to be reaching an asymptotic value at the higher brightness levels; and this is also apparent when the lack of significant difference in performance between Groups 1 and 2 is considered. For extremely high values of brightness, performance would be expected to decrease. This decrease would be due to the inability of *S* to track a target of great brilliance.

There is a significant difference between target speeds. This difference is obtained when the slow speed score has been made equivalent to the fast speed score with respect to maximum possible time on target. The following hypothesis is offered to account for the difference between target speeds. When the plane moves across the screen, *S*'s movements may be such that he overtakes the target, passes it, returns, overtakes it, etc. His movements then become less variable about the plane as it crosses the screen. If this is the case, then it follows that *S* will be on a slow target plane more consistently than on a fast target plane. This effect may be greater at the higher brightnesses, since at the lower brightness *S* has the additional problem of sighting an extremely dim plane. The significant brightness group by speed interaction may be attributed to this effect.

Both analyses on the transfer data indicate that practice at the lower brightness levels is ineffective when *S* is tested at the highest brightness level, since the transfer scores of Groups 4, 5, and 6 do not differ significantly from the first two trials of Group 1 on Day 1, and Groups 3, 4, 5, and 6 differ significantly from Groups 1 and 2. Evidently the time spent in practice is not so important as the condition under which practice takes place.

Summary

Tracking proficiency was studied as a function of target brightness. Six groups of 16 *S*s practiced for five sessions at one of six levels of target brightness. The brightness values were 697 mL, 6.97 mL, .0697 mL, .00697 mL, .0044 mL, and .0028 mL. One-half of each group practiced at one of two speeds. Two terminal transfer trials were given with the brightest target. The results show acquisition scores to fall into three brackets. The two groups practicing at the two brightest targets do not differ significantly from each other. The group practicing at the third highest brightness differs significantly from all other groups. The groups practicing at the three lowest brightnesses do not differ significantly from each other. Further analysis indicates that as target brightness increases tracking proficiency increases in a curvilinear manner. There is a significant difference between speeds in favor of the slower speed, and the brightness speed interaction is significant. The transfer trials indicate that practice at the lowest brightness level is ineffective compared to practice at the higher brightness values used.

References

1. Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.
2. Battig, W. F., Nagel, E. H., Small, A. M., Jr., Gregg, L. W., & Brogden, W. J. The effect of frequency of target intermittence on tracking. *J. exp. Psychol.*, 1954, **47**, 309–314.
3. Duncan, D. B. A significant test for differences between ranked treatments in an analysis of variance. *Virginia J. Sci.*, 1951, **2**, 171–189.
4. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
5. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
6. Melton, A. W. (Ed.) Apparatus tests. Washington, D. C.: U. S. Government Printing Office, 1947. (*AAF Aviat. Psychol. Program Res. Rep.* No. 4.)
7. Snedecor, G. W. *Statistical methods.* Ames, Ia.: Iowa State Coll. Press, 1946.
8. *Pedestal sight manipulation test. Manual on maintenance, calibration, and operation.* Davenport, Ia.: Mast Development Co., 1953.

(Received June 26, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

EFFECT OF FREQUENCY OF TARGET INTERMITTENCE UPON TRACKING[1]

W. F. BATTIG, JAMES F. VOSS, AND W. J. BROGDEN

University of Wisconsin

Battig, Gregg, Nagel, Small, and Brogden (1) have reported a study on tracking and frequency of target intermittence. The attempt was made to determine the function between proficiency of tracking and frequency of intermittent presentation of the target at .4, .7, 1.0, 2.1, 4.1, 6.0, and 8.2 cps and a steady state condition. Unfortunately, limitations of the light flash generator used to present the intermittent target conditions made it impossible to provide the desired 50/50 light-dark ratio at all frequencies. Serious deviation occurred at frequencies of .4 and .7 cps. In addition, the brightness of the target was not constant but decreased as frequency of intermittence increased. In spite of this confounding of variables, there was evidence supporting the hypothesis that proficiency of tracking varies as a function of frequency of target intermittence. There also was the possibility that, at the lowest frequency, tracking performance was poorer than at any other condition. As frequency of intermittence increased, tracking proficiency increased to a maximum at 6–8 cps, but declined at the steady state condition.

The present studies eliminate the confounding of light-dark ratio and target brightness with frequency of target intermittence. Light-dark ratio was maintained at the desired proportion of 50/50 and target brightness was .548 mL for each of the target conditions. The frequencies are 2, 4, 6, 8, 10, 12, and 14 cps. The brightness of the steady state condition (16 cps, just above fusion frequency) was also .548 mL. These values of the independent variable are different from those of the earlier study. Frequencies of 1 cps and less were not included in the present experiments. From the earlier study, it appears that tracking proficiency at these frequencies is considerably below the level of proficiency for a steady state target. If tracking proficiency is greater for some frequency of target intermittence than for the steady state target, the results of the earlier experiment indicate that it will occur within the range of 6–14 cps. The frequencies for the present experiments were selected so that determination might be made of the position of the optimum effect of intermittent presentation of the target on tracking, if such an effect does occur. Increments of 2 cps were used so that the nature of the function might be obtained by application of orthogonal polynomials to the data. Two experimental designs were used. For the first, different *Ss* performed the tracking task at each one of the eight conditions. With the task employed, however, individual differences are large and the simple factorial design, therefore, is lacking in precision. For this reason, in the second design, an 8 × 8 latin square was used

[1] This research was supported in part by the United States Air Force under Contract Number AF 33(038)-23294 monitored by the Armament Systems Personnel Research Laboratory, Air Force Personnel and Training Research Center, Lowry Air Force Base, Denver, Colorado. Reproduction in whole or in part is permitted for any purpose of the United States Government.

in order to control the effect of individual differences and increase the precision of experimentation.

Apparatus and Procedure

The Mast modification of the Pedestal Sight Manipulation Test (PSMT) was used in the present study (8). A slight change of this apparatus was necessary in order to provide for control of the frequency of intermittent target presentation by an electronic flash generator. The circuits and operation of this instrument have been described by Roush and Hamburger (5). A Sylvania Type R-1131C glow modulator tube replaced the standard bulb in the projection unit of the PSMT. Controls on the generator provide for continuous variation of frequency and light-dark ratio. Since minimum target brightness occurred at 14 cps, the brightness of the target at the other frequencies was adjusted to its value of .548 mL. This required preliminary experimentation in which the authors served as both *E* and *S*. Because of the directional character of the Scotchlite surface of the PSMT screen, it was not possible to make direct readings of target brightness with a Macbeth illuminometer. A circular plate of milk glass was illuminated by a light source, the intensity of which could be varied. It was placed adjacent to the target and each of the three *S*s adjusted the brightness of the milk glass surface until it equaled the brightness of the target for each of the frequencies and for the steady state condition. Since there were no significant differences between *S*s, the data were pooled and were analyzed by analysis of variance and by the Duncan range test (2). No statistically significant differences were obtained between the brightness match for 14 cps and the matches for 8, 10, and 12 cps. The brightness of the steady state condition (16 cps) was found to be significantly less than that for 14 cps, and those for 2, 4, and 6 cps were found to be brighter. It was possible to increase the duration of the light phase so that the brightness of the steady state target at 16 cps was equal to the brightness of the target at 14 cps. Macbeth readings of the brightness of the matches made it possible to designate these target brightnesses in mL and to compute the density of filters needed to provide a brightness of .548 mL. Filters with values of .2, .15, and .1 log density were placed in the projection system to reduce the brightness of the target to .548 mL at 2, 4, and 6 cps, respectively.

The scoring area was set at 25 mils (1.5 in.) for both azimuth and elevation. The range band was set at 20 mils (1.0 in.). Standard Electric clocks were used to record time-on-target scores in seconds for azimuth (A), elevation (E), and range (R), for azimuth and elevation (AE), and for azimuth, elevation, and range (AER). Thus five separate measures of tracking proficiency were obtained.

The 80 *S*s were male undergraduate students at the University of Wisconsin. In Exp. 1, each *S* was assigned at random to one of eight groups, a group representing a single condition of target presentation. The eight groups were built up simultaneously. Each *S* received eight trials in one practice session of approximately ½-hr. duration. A trial is made up of a standard sequence of eight flights of the target plane across the screen. The intertrial interval was 30 sec., except that between the fourth and fifth trials the interval was 2 min.

Experiment 2 used an 8×8 latin square. Two *S*s from each of the eight groups of the first experimental design were selected and assigned at random to each of the eight sequences of target conditions. Two trials were given for each condition and the interval between them was 30 sec. The interval between pairs of trials, i.e., different target conditions, was 2 min. The single practice session lasted for approximately 1 hr.

For Exp. 1, the instructions described the pedestal sight and its controls, and *E* demonstrated to *S* how to track and frame. Scoring was explained and *S* was informed that the target might go on and off intermittently. He was instructed to continue movement of the sight and range knob in anticipation of the course of the target. The *S* was given practice in manipulating the sight and, after the kneeling platform was adjusted and any questions answered, the experiment was begun. The *E* left the experimental room to operate the control apparatus. For Exp. 2, *S* was reminded of the previous instructions and was told of the new experimental conditions involving different frequencies of target intermittence.

In general, *S*s were tested during the mornings between 8:00 A.M. and 1:00 P.M., and in the evenings between 6:30 P.M. and 10:30 P.M. About half the *S*s in each experiment were run at each of these two times. The three *E*s were not orthogonally represented in either of the two experimental designs and different numbers of *S*s were run by each one depending upon schedule of other duties. Calibration checks of the azimuth and elevation scoring area and of the range band were made before each morning session and each evening session of the experiment.

Results

The group means of time-on-target tracking performance in terms of A, E, R, AE, and AER were computed for each of the eight target conditions for each experiment and graphic plots were made. Since the form of the functions for each of the five measures is very similar, no further analysis of the A, E, and R measures was conducted. Figure 1 presents a plot of the group mean AE and AER performance as a function of intermittent presentation of the target for both experiments. Table 1 gives a summary of the results of analysis of variance of the AE and AER data of Exp. 1 and 2. The assumption of homogeneity of variance for the data of Exp. 1 was verified by Bartlett's test (3). Only the F ratio for the AE measures, however, is significant at the 5% level of confidence. For Exp. 2, the procedure for analysis of variance follows Grant (4). Target conditions is a significant source of variation for both the AE and AER data. The only other significant source of variation is that for *Ss* within sequences. This occurs for both the AE and AER data and indicates significant individual differences between *Ss*.

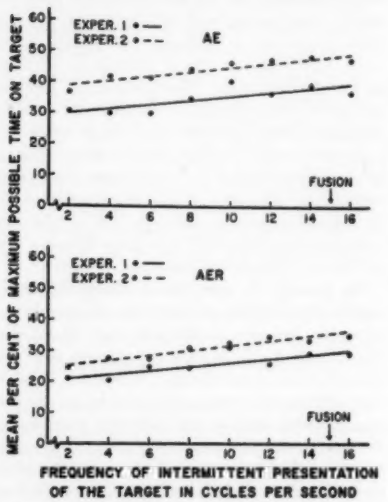

Fig. 1. Tracking proficiency as a function of frequency of intermittent presentation of the target.

The results of analysis of variance show that there are significant differences in mean AE performance as a function of frequency of intermittent presentation of the target for both Exp. 1 and 2, and for mean AER performance for Exp. 2. More precise information about these functions may be obtained by the use of Duncan's range test (2) and of orthogonal polynomials (6). The results of the Duncan range test need not be given in detail. They indicate significant differences between mean performance as the frequency of intermittent presentation of the target increases. In no case where there is a decrement in performance with an increase in frequency (including the steady state condition) is the difference statistically significant. For each of the four sets of data, including the AER data of Exp. 1, the application of orthogonal polynomials reveals a significant linear component. In no case are any higher order components of significance.

Since the above analyses establish the linearity of the obtained functions, straight line regression equations were computed by the method of least squares. These equations are: $y = .71x + 28.59$ for AE, Exp. 1; $y = .77x + 37.41$ for AE, Exp. 2; $y = .62x + 20.10$ for AER, Exp. 1; and $y = .78x + 23.73$ for AER, Exp. 2. These equations have been plotted in Fig. 1.

Discussion

The results of the present experiments provide no reliable evidence that tracking proficiency is at an optimum level at

TABLE 1

SUMMARY OF RESULTS OF ANALYSES OF VARIANCE

Source	*df*	AE Mean Square	*F*	AER Mean Square	*F*
		Experiment 1			
1. Between groups	7	3419.2	3.92*	2570.6	1.47
2. Within groups	72	872.6		1746.9	
		Experiment 2			
1. Target conditions	7	74.8	19.68*	72.6	15.13*
2. Ordinal position	7	3.9	1.03	6.9	1.44
3. Sequence	7	136.1	1.34**	296.2	0.68**
4. Square uniqueness	42	4.7	1.24	6.3	1.31
5. *Ss* within sequence	8	101.9	26.82*	436.7	9.10*
6. Residual error	56	3.8		4.8	

* Significant at the 5% level of confidence or better.
** Denominator of *F* ratio is mean square for *Ss* within sequence.

some frequency of intermittent presentation of the target below fusion frequency. Since the decrement in tracking proficiency observed by Battig, Gregg, Nagel, Small, and Brogden (1) between 8.2 cps and the steady state condition (16.1 cps) could have been due to the decrease in brightness, the results of the present experiment favor this interpretation of their results. Further support of this interpretation comes from a recent study by Voss (7) that shows tracking proficiency to decrease with decrease in the brightness of the target. His results, however, indicate that a substantial change in brightness (greater than 1 log mL) is required to produce a significant difference in tracking performance. Since the difference in brightness at 8.2 cps and the steady state condition was less than 1 log unit in the experiment by Battig, Gregg, Nagel, Small, and Brogden, there is the possibility that the increment in tracking performance at 8.2 cps over the steady state condition may be a function of both the intermittent presentation of the target and the greater target brightness. The unreliable numerical superiority of tracking performance at some frequency of intermittent presentation of the target over that of the steady state target for all conditions of the present experiment except the AER measures of Exp. 2 gives some support to this latter interpretation.

The increase in tracking proficiency as a function of increase in frequency of intermittent presentation of the target found by Battig, Gregg, Nagel, Small, and Brogden has been confirmed by the present experiments for the range between 2 and 14 cps. With extension to include the steady state condition, the least square regression equations provide an adequate quantitative expression of the functional relation of tracking proficiency and frequency of intermittent presentation of the target. Slope appears to be very nearly the same, no matter what the other conditions. The *y* intercept, however, is smaller for AER than for AE, reflecting no doubt the relatively greater difficulty of AER performance. The *y* intercept is less for Exp. 1 than for Exp. 2, and this undoubtedly reflects a practice effect, since the *Ss* of Exp. 2 had previously performed in Exp. 1. These results suggest that the slope of the function is independent of the particular measure of tracking and of amount of practice, but that each is an important parameter in determining the *y* intercept.

It should be noted that regression equations of Exp. 2 provide for a better fit of the empirical data than do the

regression equations fitted to the data of Exp. 1. This may be due to the greater amount of practice had by the *S*s in Exp. 2. It may also be a function of the experimental design. The latin square makes it possible to control individual differences and provides greater efficiency in experimentation in that a smaller number of *S*s may be used. Although the present study is not a precise test, the results do support the theoretical greater efficiency of the latin-square design (4).

Finally, comment should be made of the significance of the results of the present experiments in relation to tracking behavior during flight of the target. For the conditions of intermittent presentation, the target is observed by *S* for only half the time of its flight, no matter what the frequency of intermittence. As frequency is increased from 2 to 14 cps, the duration of each presentation of the target and of each dark phase is decreased. Thus, as frequency increases, the course of the target is better defined and *S* is better able to learn the target courses and move the sight in anticipation of movement of the target. More frequent information is also available to *S* about the accuracy of his tracking. He, therefore, is better able to reduce his error and increase his time-on-target score. It is these features of intermittent presentation of the target that probably are responsible for the obtained functions between tracking proficiency and frequency of intermittent presentation of the target.

Summary

Measures of proficiency of tracking on the Pedestal Sight Manipulation Test were obtained at frequencies of intermittent target presentation of 2, 4, 6, 8, 10, 12, and 14 cps and a steady state condition above fusion (16 cps). The brightness of the target was held constant for all conditions at .548 mL and the light-dark ratio for each of the frequencies was 50/50. Two experiments were performed. In one, ten *S*s each received eight trials of practice for one of the eight conditions of target presentation. In the other, two *S*s were selected at random from each of the eight groups of the first experiment and these 16 *S*s received two trials of practice on each of the eight target conditions as determined by an 8×8 latin square.

Although measures of tracking proficiency were obtained for azimuth (A), elevation (E), range (R), azimuth and elevation (AE), and azimuth, elevation, and range (AER), because of the similarity of results over these separate measures, only the AE and AER data were analyzed completely and presented. Analysis of variance demonstrated a significant effect of frequency of intermittent presentation of the target on tracking. Further statistical analysis revealed the relations to be linear, tracking proficiency increasing with increase in frequency of intermittent presentation of the target. Least-square regression equations were fitted separately to the AE and AER data for the two experiments. Although slope is essentially the same for the several conditions, the *y* intercept differs as a function of the measure of tracking (AE vs. AER) and of amount of practice (Exp. 1 vs. Exp. 2). The results are discussed especially in relation to those of a prior experiment by Battig, Gregg, Nagel, Small, and Brogden.

References

1. Battig, W. F., Gregg, L. W., Nagel, E. H., Small, A. M., Jr., & Brogden, W. J. Tracking and frequency of target intermittence. *J. exp. Psychol.*, 1954, **47**, 309–314.
2. Duncan, D. B. A significance test for differences between ranked treatments in an analysis of variance. *Virginia J. Sci.*, 1951, **2**, 171–189.
3. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
4. Grant, D. A. The latin square principle in the design and analysis of psychological experiments. *Psychol. Bull.*, 1948, **45**, 427–442.
5. Roush, R. G., & Hamburger, F., Jr. A light flash generator. *Electronics*, 1948, **21**, 100–102.
6. Snedecor, G. W. *Statistical methods.* Ames, Ia.: Iowa State Coll. Press, 1946.
7. Voss, J. F. The effect of target brightness and target speed upon tracking proficiency. *J. exp. Psychol.*, 1955, **49**, 237–247.
8. *Pedestal sight manipulation test. Manual on maintenance, calibration, and operation.* Davenport, Ia.: Mast Development Co., 1953.

(Received July 1, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

THE FREQUENCY RESPONSE OF SKILLED SUBJECTS IN A PURSUIT TRACKING TASK [1]

MERRILL NOBLE,[2] PAUL M. FITTS, AND CLAUDE E. WARREN

The Ohio State University

Frequency response functions are widely used in specifying the functional characteristics of physical systems. Such functions indicate the amplitude and phase relations between input and output for sinusoidal input signals of different frequencies. Human response capabilities can be specified in a similar manner. By systematically increasing the frequency of a sinusoidal stimulus motion *E* can study factors involved in the breakdown of the ability to perform a continuous tracking task as the temporal relation between stimulus and response becomes more and more critical. The use of a sinusoidal stimulus motion appears to rest on a sound physiological basis since Fenn has pointed out that in rapid oscillatory arm motion "evidently the central nervous system 'deals out' excitation to the appropriate muscles according to a sine wave pattern . . ." (**6**, p. 172). It should be emphasized, however, that a sinusoidal target motion is highly coherent and thus the human frequency response obtained with this type of stimulus reflects *S*'s ability to cope with a predictable time-varying function.

In most physical systems the output amplitude decreases after some critical input frequency is reached. Maximum output amplitude often occurs at some intermediate frequency, known as the resonant frequency of the system. Input and output phase relations, however, vary widely from one physical system to another.

The results of several studies (**8, 9, 10, 12, 13**), in which only a few selected frequencies of target motion are used, indicate that human tracking performance tends to become worse as the frequency or average velocity of the response has to be increased. Graphic error records of unskilled *Ss* for 30- and 90-cpm input frequencies obtained in earlier studies with the present apparatus (**8, 13**) indicate that as frequency is increased response movements become smoother within a single cycle but show greater variability from cycle to cycle, and exhibit higher peak accelerations. With increased practice, responses at all frequencies exhibit smaller peak accelerations, are smoother throughout a cycle, and are less variable from cycle to cycle. Ellson and Gray (**4**) and Ellson and Wheeler (**5**), employing a pursuit tracking task, frequencies ranging from 30 to 240 cpm, and moderately skilled *Ss*, found no significant reduction in amplitude of graphic records at higher frequencies; instead, the records indicated a loss of stimulus-response synchronization at higher frequencies. After loss of synchronization most *Ss* maintained an output frequency that was somewhat higher than that of the input.

[1] This research was supported in part by the United States Air Force under Contract No. AF 33(038)-10528 with The Ohio State University Research Foundation, monitored by the Air Force Personnel and Training Research Center. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government. The aid of Ilse Webb and James Polidora in collecting the data and Dr. Philip Siegmann in performing the statistical analyses is acknowledged.

[2] Now at Kansas State College.

Muscle action potential records revealed relatively lower amplitude at certain frequency ranges indicating a possible region of resonance. No consistent phase shift pattern was discovered.

In the present investigation quantitative measures of tracking error in addition to sample graphic records were taken for a range of target frequencies between 10 and 240 cpm. Highly practiced *S*s were employed in order to provide a stable estimate of the frequency response which can be attained with extensive practice. Over-all tracking efficiency was evaluated by means of the quantitative scores; phase and amplitude changes were analyzed by use of the graphic records.

Method

Apparatus

An electronic tracking apparatus (**14**) was used. Essentially, it consists of (*a*) a problem generator which provides a continuous target signal, (*b*) a 5-in. cathode ray tube (CRT) upon which *S* can observe a target line and his own response or cursor line, (*c*) an arm control, (*d*) scoring and recording units, and (*e*) a programming unit.

The display.—The target appeared as a line 18 mm. long and .5 mm. wide which moved horizontally across the center of the CRT display. The cursor had the same dimensions and was displaced 10 mm. below the target. Thus, when the target and cursor were superimposed they formed a single line 28 mm. long with an 8-mm. region of overlap. A short-persistence P-11 phosphor CRT was used. The surface of the CRT was approximately 28 in. from *S*'s eyes, 10° below eye level, and perpendicular to the line of regard. The booth in which the display was located was dimly illuminated, and specular reflection from the CRT was eliminated by means of polaroid filters. Target and cursor lines were painted on the display alternately by a 200-cps time-sharing switch. Because of the high rate of time sharing, no stroboscopic effect was apparent except at very high rates of cursor movement. At extreme rates the cursor appeared to move in discrete steps.

The control.—The *S* sat in a chair which could be adjusted for height. The elbow of his right arm rested upon the pivot point of a horizontal lever, and his fingers grasped an adjustable knob on the end of the lever. Attached to the control was a potentiometer. The voltage across this potentiometer determined the position of the cursor on the CRT.

Tracking problem.—The output of the problem generator caused the target line to move back and forth in the horizontal plane on the CRT. The time course of this motion was a simple sine wave. The amplitude of target movement was .5-in. (1.02° of visual angle) displacement of the target on either side of the center of the CRT. The amplitude of control movement required to move the cursor a corresponding amount was 10° to either side of the center position of the control, measured in degrees of rotation around the pivot point below the elbow. The *S*s were instructed to move the arm control so as to keep the cursor line superimposed on the target line. Error is defined as the difference between the output voltage of the target generator and the voltage across *S*'s arm control.

Simple sinusoidal target courses of ten different frequencies were studied—10, 20, 40, 60, 90, 120, 150, 180, 210, and 240 cpm. Input amplitude remained constant for all frequencies.

Subjects

Three highly skilled right-handed male *S*s were employed. Each *S* had served in previous tracking studies and had received between 500 and 1000 1-min. trials with similar target courses (at frequencies ranging from 30 to 90 cpm). They received hourly wages.

Procedure

Each *S* received 14 1-min. preliminary practice trials, separated by 1-min. rest periods, on each of the ten frequencies studied. These 140 preliminary practice trials were given during seven 1-hr. sessions, each separated by at least 24 hr. During each 1-hr session two successive trials were given on each of the ten frequencies studied; otherwise, the order of frequencies was determined separately for each *S* by a table of random numbers. A different random order was followed by each *S* on each session.

Following the practice sessions, each *S* received a total of 24 test trials on each of the ten frequencies studied, or a total of 240 test trials. Each trial was 66 sec. in duration, but the first 6 sec. were treated as a warm-up period and were not scored. Test trials were given during 12 1-hr. sessions, each separated by at least 24 hr. The procedure, including determination of the sequence of presentation of

different frequencies, corresponded to that used during the practice trials. At the end of this series of test sessions, additional trials were given during which simultaneous graphic records of the target course, control movement, and error were obtained.

Quantitative Performance Criteria

Two quantitative measures of tracking error were obtained: (*a*) root-mean-squared error (error *RMS*), and (*b*) time-on-target within a tolerance of 5% of the maximum plus or minus problem voltage (this is equivalent to a ±.025-in. misalignment between target line and cursor line on the tracking display). Error *RMS* was obtained by taking the square root of a meter reading obtained from circuits which accumulated the squared error voltage during each trial. Previous studies employing frequencies not exceeding 90 cpm have shown that the error *RMS* and the 5% tolerance time-on-target scores give a highly correlated but not identical description of performance (**8, 13**).

Results

Objective scores.—Error *RMS* scores pooled for the first six and the last six of the series of 12 trials are shown in Fig. 1 as a function of frequency for each *S* separately. The dotted line at an ordinate value of 29.3 represents the *RMS* value of the input voltage from the problem generator. An error *RMS* score in excess of this value represents performance worse than would be achieved if *S* centered his control and made no effort to track the target during a trial. This score does not represent complete breakdown, however, since if *S* should continue to move the control in approximately sinusoidal motion but with a frequency different from that of the

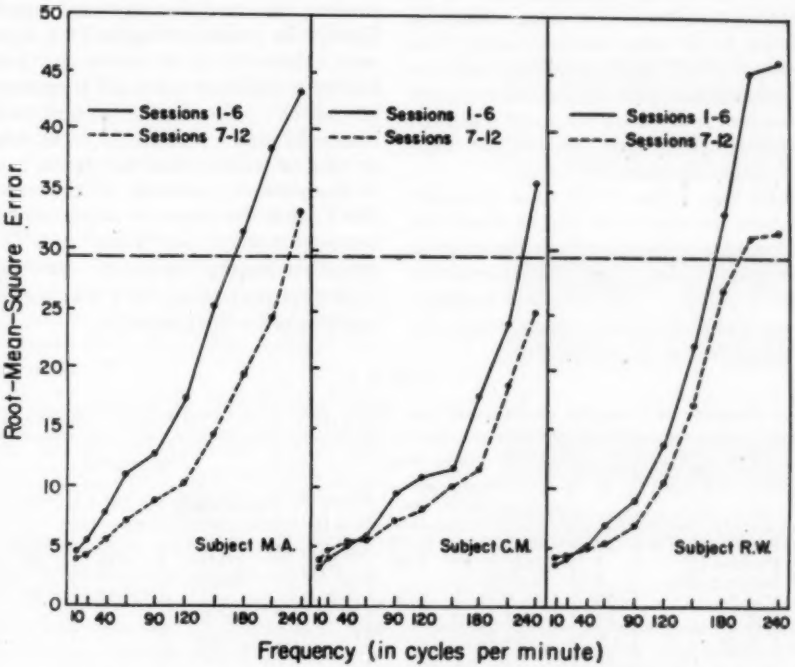

FIG. 1. Error *RMS* scores of three *S*s as a function of target frequency. The dotted line represents the *RMS* value of the target.

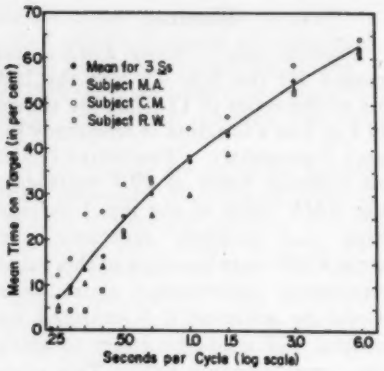

FIG. 2. Time-on-target score as a function of the duration of a stimulus cycle. Data are for Trials 7–12.

target, his average *RMS* score would be considerably greater than 29.3. It is obviously inefficient for *S* to attempt to track the target when by doing so his error becomes larger than it would be if he did nothing, and it is interesting to note that with continued practice individual error curves reach approximately this value at the highest input frequencies.

In Fig. 2 the mean time-on-target scores for the same set of trials are shown for each *S* as a function of cycle duration, using a logarithmic scale for these values. The solid line in Fig. 2 is a visual fit of the mean scores for the three *Ss* combined. While the data of Fig. 2 are fitted with a negatively accelerated curve, a straight line would fit the data almost as well. Furthermore, the time-on-target scores are attenuated slightly at the lower end of the cycle duration scale, owing to a slight lag in the scoring clocks which begins to exert an appreciable attenuating effect on scores where a large number of very short-duration on-target periods must be accumulated. Fortunately, however, the error *RMS* score is free of this artifact.

It is apparent from an examination of Fig. 1 and 2 that no point of marked discontinuity in performance occurs with increase in frequency. The smooth error *RMS* curves, of course, are in part a result of the pooling of data for repeated trials. Except for chance irregularities, however, there is no evidence in these curves to indicate a critical frequency at which performance deteriorates markedly and consistently from trial to trial or from person to person.

Analyses of variance of the error *RMS* and the time-on-target scores are summarized in Table 1. The variation among the means resulting from changes in frequency was highly significant for both scores.

TABLE 1

SUMMARY OF ANALYSES OF VARIANCE FOR ERROR *RMS* AND TIME-ON-TARGET SCORES

Source of Variance	*df*	Error *RMS*			Time on Target		
		Mean Square	*F**	*p*	Mean Square	*F**	*p*
Frequency (F)	9	14087.29	32.75	.001	1328190	95.51	.001
Practice (P)	3	1956.31	5.85	.05	234544	6.45	.05
Subjects (*Ss*)	2	3195.97			253752		
F × P	27	179.56	5.08	.001	6113	1.57	.10
F × *Ss*	18	430.20			13906		
P × *Ss*	6	334.26			36369		
F × P × *Ss*	54	35.36			3884		

* The interaction F × *Ss* was used as the error term to test the effect of Frequency; P × *Ss* to test the effect of Practice; and F × P × *Ss* to test the effect of Frequency by Practice. Variation among *Ss* was not evaluated.

Even though skilled *Ss* were employed, some learning was apparent, and it was decided to evaluate the effect of practice during the 12 sessions. With four successive blocks of three trials each for this assessment, a significant improvement in performance ($p < .05$) as a function of practice was found for both scores. The interaction between frequency and practice was statistically significant for the error *RMS* scores but not for the time-on-target scores, practice leading to relatively greater improvement at the higher frequencies.

Graphic records.—Sample graphic records of tracking error are shown in Fig. 3. Sample simultaneous records of the stimulus and response motion for one *S* are shown in Fig. 4. The increase in error amplitude at the higher frequencies is primarily a result of lack of synchronization of *S*'s output frequency with the input frequency. This interpretation, which was reported earlier by Ellson and associates (**4, 5**), is verified by the increasing presence of the target frequency in the error records (see especially the records at 150 cpm), and by the relatively common occurrence of beats at very high frequencies. The number of beats per minute in each case corresponds to the difference between the input and output frequencies.

There were large individual differences both in the frequency at which *Ss* lost synchronization and in the frequency which characterized their output following loss of synchronization. As an illustration, one *S* (M. A.) began to lose synchronization for input frequencies above 180 cpm. At 210 cpm his mean output frequency was 232 cpm, and at 240 cpm it was 250 cpm. Another *S* (R. W.) began to lose synchronization at 150 cpm. At an input of 180 cpm his mean output was 160 cpm, at 210 cpm it was 155 cpm, and at 240 cpm it was 160 cpm. In all graphic records the output frequencies were fairly consistent for the same *S* from sample to sample. As an illustration, four independent determinations of the output rate of

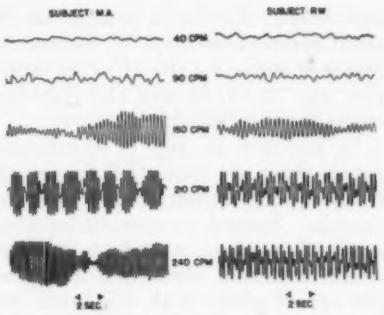

FIG. 3. Sample graphic records of tracking error for two *Ss*.

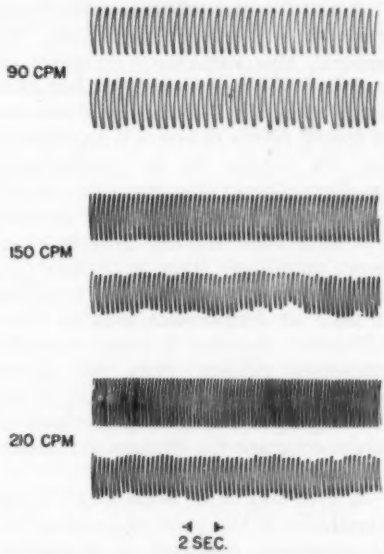

FIG. 4. Sample graphic records of stimulus and response motion taken simultaneously for *S* R. W. At each frequency the top record is the stimulus, the bottom record the response motion.

one *S* (M. A.) for a sample of 50 input cycles showed 56, 59, 56, and 54 response cycles for the 210 cpm input rate, and 52, 51, 54, and 51.5 response cycles for the 240 cpm input rate.

In addition to loss of synchronization there are four other sources of tracking error apparent in the graphic records. Second, at input frequencies lower than those at which *S* loses synchronization, his output tends to get out of phase with the input for short periods of time. At the lower frequencies this loss of phase relation usually lasts for only ½ cycle, while in the middle range of frequencies *S* may stay out of phase for 2 or 3 cycles before a relatively good correction of his phase error is made. In other words, *S* apparently requires a relatively constant time interval to observe and correct phase errors irrespective of input frequency.

A third souree of error arises from variability of output amplitude. In general, this variability increases as problem frequency increases, but here again there are individual differences. A fourth factor in errors is a tendency for *S*'s output to be consistently smaller or larger than the input amplitude. One *S* tended to have an output amplitude slightly greater than input amplitude (that is, to make all movements greater than the specified ± 10°) at frequencies greater than 120 cpm. Another *S*, however, made responses smaller than the input amplitude, his output amplitude decreasing up to 150 cpm and remaining fairly constant for frequencies above 150 cpm.

A fifth source of error arises from variability in the point of termination of successive extensor or successive flexor movements. Variability in this aspect of the movements is particularly pronounced in the records of the *S* (R. W.) shown in Fig. 4. This type of variability tends to increase steadily as a function of frequency, at least up to input frequencies of 210 cpm.

Discussion

It appears from the present results that *S*s can follow an input frequency up to about 2 cps and keep their error *RMS* score less than half the value of the signal itself. Performance does not fall to the level of the problem signal until around 4 cps. This is a considerably higher upper frequency than the limit of .5 cps estimated by Birmingham and Taylor (1) for a compensatory display and irregular target motion. It must be recalled, however, that the present data were obtained under conditions favorable to good tracking—highly predictable target motion and a pursuit display. The upper limit of frequency is actually slightly above the maximum frequency (about 3 cps) observed in a task in which *S*s were required to control the amplitude of reciprocal arm movements (7). The human frequency response undoubtedly will exhibit much earlier loss of efficiency for less predictable target courses and for tracking displays that do not permit direct observation of the pattern of target motion.

The present results reveal several interesting differences between the human frequency response and that of electronic or mechanical devices. No consistent shifts in the mean phase relations of input and output occurred short of the point at which synchronization of stimulus and response was lost. No consistent indications of a resonant frequency were found, nor was there any consistent decrease in response amplitudes at the higher frequencies. Performance at high frequencies was not consistently accompanied by reduction in amplitude, as is true of most physical systems, but evidenced periods of loss of stimulus-response synchronization and increased variability in the relation between input and output. These variations appear to increase steadily in importance prior to the loss of stimulus-response synchronization.

The model of human tracking performance that appears to fit the preceding experimental facts most adequately is one that involves two assumptions: (*a*) a basic intermittency of data processing by the human, and (*b*) a predictive capacity by means of which the human, operating intermittently on sampled data, adjusts his output to correspond approximately to a continuous input. It is assumed that in a continuous task in which the temporal changes of the stimulus are highly predictable man's output is relatively continuous and is based primarily on a prediction of the future stimulus pattern. It is assumed further that man compares a sample segment of the stimulus pattern with a sample of his own response pattern, and uses this information to regulate, at intervals, his own continuous output in order to bring it into closer correspondence with the input pattern. These adjustments may take the form of periodic shifts in the points of termination of flexor or extensor movements, variations in the mean and range of a series of movements, or adjustments of the frequency and phase relations of successive responses. In other words, it is proposed that in a task that requires him to make his response output correspond in time and space to a continuously and rapidly changing stimulus, *S* does not respond directly to the stimulus at all—instead, he puts out a continuous predicted response, observes the relation between his output and the stimulus pattern, and at periodic intervals makes necessary changes in his predictions. If *S*'s sampling period is independent of frequency this would explain why it is that tracking error increases so rapidly at high frequencies, for if the absolute magnitude of the sampling interval is independent of frequency, then phase error will increase with frequency and we have seen that error at higher frequencies is predominantly determined by timing or phase error.

Fenn (6) has concluded that the maximum rate at which simple reciprocal movements can be made in the absence of any external stimulus is determined not by inability of the motor system to produce the forces required to make rapid movements, but by a breakdown of the control of the temporal relations required to maintain a rapid alternation between reciprocally innervated muscle groups. This breakdown occurs at or before 10 cps. The loss of tracking proficiency at higher frequencies of target motion probably arises in a similar manner, due not to *S*'s inability to discriminate and/or produce the required amplitudes of movement, but to his inability at these frequencies to perceive the temporal relations between stimulus and response patterns and to control the temporal organization of his own motor behavior.

Summary

The efficiency of highly skilled *S*s in tracking a target which moves in a simple harmonic motion pattern, as measured by root-mean-squared error and by time on target within 5% of the target amplitude, was found to decrease steadily as input frequency was increased. Examination of objective scores and analysis of graphic records of the problem input, *S*'s output, and tracking error indicate that five types of changes in motor behavior contribute to the decrease in performance proficiency: (*a*) variability in phase relations between stimulus and response, (*b*) variability of output amplitude, (*c*) variability in the point of termination of successive flexor or of extensor movements, (*d*) a constant error in matching response amplitude to the average input amplitude, and (*e*) loss of synchronization between output and input frequencies, i.e., periods during which the output frequency was independent of the input frequency. During short periods of loss of synchronization the output frequency might be either higher or lower than the input frequency.

The last of the above factors operates only at the higher input frequencies. The other four factors tend to increase steadily in importance as input frequencies increase, although variability in the point of termination of successive extensor and flexor movements tends to reach a maximum at an 150–180 cpm input frequency range. The above factors were present for each *S*, although there were marked individual differences in the relative influence of the several factors at the various frequencies.

These changes in performance are interpreted in relation to the hypothesis that man puts out a patterned response which he predicts will match the stimulus pattern, observes the

outcome during a sampling period, and intermittently changes his output pattern.

REFERENCES

1. BIRMINGHAM, H.P., & TAYLOR, F. V. A human engineering approach to the design of man-operated continuous control systems. *U. S. Naval Res. Lab. Rep.*, 1954, No. 4333.
2. CRAIK, K. J. W. Theory of the human operator in control systems: I. The operator as an engineering system. *Brit. J. Psychol.*, 1947, **38**, 56–61.
3. CRAIK, K. J. W. Theory of the human operator in control systems: II. Man as an element in control systems. *Brit. J. Psychol.*, 1948, **38**, 142–148.
4. ELLSON, D. G., & GRAY, F. Frequency responses of human operators following a sine wave input. *USAF, Air Materiel Command, Memo. Rep.*, 1948, No. MCR-EXD-694-2N.
5. ELLSON, D. G., & WHEELER, L., JR. Resonance in the human operator. *USAF, WADC Tech. Rep.*, 1951, No. 5834.
6. FENN, W. O. The mechanics of muscular contraction in man. *J. appl. Phys.*, 1938, **9**, 165–177.
7. FITTS, P. M. The information capacity of the human motor system in controlling the amplitude of movement. *J. exp. Psychol.*, 1954, **47**, 381–391.
8. FITTS, P. M., MARLOWE, E., & NOBLE, M. E. The interrelations of task variables in continuous pursuit tasks: I. Visual display scale, arm control scale and target frequency in pursuit tracking. *USAF, Hum. Resour. Res. Cen., Res. Bull.*, 1953, No. 53-34.
9. HARTMAN, B. O., & FITTS, P. M. Relation of stimulus and response amplitude to tracking performance. *J. exp. Psychol.*, 1955, **49**, 82–92.
10. HELSON, H. Design of equipment and optimal human operation. *Amer. J. Psychol.*, 1949, **62**, 473–497.
11. HICK, W. E., & BATES, J. A. V. The human operator of control mechanisms. London: Ministry of Supply, 1950.
12. MCGUIRE, J. C. The precision of tracking in two dimensions as a function of target velocity and aiming tolerance limits. Unpublished doctor's dissertation, Washington Univ., 1954.
13. NOBLE, M. E., FITTS, P. M., & MARLOWE, E. The interrelations of task variables in continuous pursuit tasks: II. Visual display scale, arm control scale, and target frequency in compensatory tracking. *USAF, Hum. Resour. Res. Cen., Res. Bull.*, 1953, No. 53-55.
14. WARREN, C. E., FONTAINE, A. B., & CLARK, J. R. A two-dimensional electronic pursuit apparatus. *USAF, Hum. Resour. Res. Cen., Res. Bull.*, 1952, No. 52-26.

(Received for early publication December 28, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

THE INFLUENCE OF NONREINFORCEMENT OF A COMPONENT OF A COMPLEX STIMULUS ON RESISTANCE TO EXTINCTION OF THE COMPLEX ITSELF

DELOS D. WICKENS AND JOHN D. SNIDE

The Ohio State University

In a study investigating the gradient of primary stimulus generalization to tones (3) the writers established conditioned galvanic reactions to one of three tones, each of which was separated by 25 j.n.d. During the acquisition trials, the control groups received 16 presentations of a particular tone accompanied by reinforcement (shock). The procedure for the experimental groups was identical, except that in addition to the 16 presentations of the tone, a click was randomly presented 12 times without reinforcement. Each subgroup of the total experiment then received eight extinction trials on one or another of the tones. Magnitude of response was measured in terms of log micromhos.

In the control groups a gradient of generalization was found, and it seemed to be generated by the extinction process, for it became steeper with progressive extinction trials. In the experimental groups no gradient was discovered; and the failure to obtain a gradient for the experimental curves seemed attributable to the high resistance to extinction of the groups tested to the generalized tones. In interpreting these results, we suggested that the click might be thought of as being a component of the tone since both stimuli share the common quality of being auditory events which come on suddenly to break the silence in the experimental situation. If this be the case, then one might further postulate that the experimental groups were undergoing a procedure which is similar to a partial reinforcement procedure and hence a procedure which produces responses that are highly resistant to extinction. A more general statement of our hypothesis would be: If a complex stimulus is presented along with reinforcement and an element of the complex is presented at some other time without reinforcement, then this latter action increases the resistance to extinction of the complex stimulus itself. It is the purpose of the present experiment to test this hypothesis more directly.

The general procedure in this experiment consisted of employing as the CS a light and a tone which came on simultaneously and terminated simultaneously. The control and experimental groups were conditioned and extinguished to this stimulus complex; for the experimentals one or other aspect (the light alone or tone alone) of the total complex was occasionally presented during the training without reinforcement. Our hypothesis would predict that the experimentals would show greater resistance to extinction than the controls.

METHOD

Apparatus.—The stimulus complex consisted of a 1000-cps tone at 40 db above threshold produced by a Jackson oscillator and delivered through earphones, and a 2-w. red light located directly in front of, at eye level, and about 5 ft. from *S*'s face. The GSR circuit employed was that described by Haggard and Gerbrands (1). Its major units consisted of a General Electric Photoelectric Recorder and an RCA Volt

Ohmyst. Time relations between CS and UCS were controlled by an electronic timing unit. The shock was produced by a DC current of 100 v. with a 120,000-ohm resistance in series with *S*.

The *E* and *S* sat in adjacent rooms. The experimental room was soundproofed with walls and ceilings painted white. It was illuminated by a 7½-w. light located overhead.

Subjects.—The 42 *S*s who served in the experiment were volunteers from the elementary psychology classes at The Ohio State University.

Procedure.—Three groups were employed with 14 *S*s in each group. Group A served as a control. The procedure in the learning trials for this group consisted of 16 simultaneous presentations of the tone and red light followed by an electric shock. The duration of this stimulus complex was 500 msec. After the tone and light had been on for 300 msec., the shock was administered for 200 msec., at the end of which all three terminated simultaneously. The test trials consisted of eight presentations of tone and light together without reinforcement. The interstimulus interval varied randomly from 20 to 60 sec. in a prearranged order identical for all *S*s.

Two experimental groups were employed, Groups B and C. The procedure for Group B was essentially similar to that for Group A. However, the *tone* alone was administered 12 times during the learning trials in addition to the 16 coincident presentations of the tone and light. The single tone was never followed by shock.

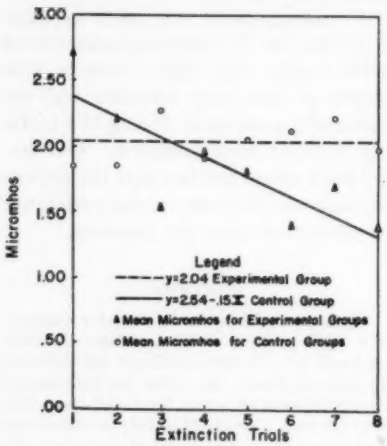

FIG. 1. Extinction curves for the experimental and control groups.

For Group C, the *red light* alone was presented 12 times during the learning trials without reinforcement. The extinction trials for Groups B and C were identical with those for Group A, that is, they were extinguished to the complex of tone plus light. The unreinforced tone or light was presented in a random sequence between the occurrence of the reinforced complex. A sequence of two consecutive stimulations of the tone or light alone was never employed.

RESULTS AND DISCUSSION

Graphs of the data are presented in Fig. 1. The separate points in the figure give the mean values in micromhos for the particular extinction trial.

Since Groups B and C did not differ from each other, they have been combined. This failure for the two experimental groups to show a difference is to be expected. Conceptually they were treated alike, and before the experiment was begun an effort was made to equate the psychological intensity of the light and the tone.

There was some tendency for the control group to show a higher magnitude of response on the initial extinction trial, but the trend is not significant.

The curves are the best fit for the data, and they are both straight lines. The equation for Group A is

$$P(x) = 2.54 - .15x,$$

while the equation for the combined Groups B and C is

$$y = 2.04.$$

There is no significant difference between the displacement of the two curves, but the slope terms differ from each other at the 5% level of confidence in a two-tailed test (2).

Thus the data of the experiment indicate that the experimental groups (B and C) show greater resistance to extinction than the control group.

The results of the experiment support the hypothesis which instigated the

experiment. The occurrence without reinforcement of a component of a stimulus complex which itself is consistently reinforced increases the resistance to extinction of the complex. If this state of affairs is generally true, then the operation of partial reinforcement is considerably more subtle and considerably more pervasive than at least superficially seems to be the case. It will be noted, for example, that the process of discrimination training partakes of this character. One might, therefore, expect that the positive stimulus of the discrimination pair obtains a higher resistance to extinction than would have been the case if it alone had been presented for the same number of reinforcements.

Summary

The present experiment was a more direct test of a hypothesis which was advanced to account for results obtained in a previous study of stimulus generalization. It was hypothesized that if a complex stimulus is presented along with reinforcement and an element of the complex is presented at some other time without reinforcement, then this later action increases the resistance to extinction of the complex stimulus itself.

To this end, the control *S*s were conditioned to a complex stimulus of a light and tone for 16 trials, with shock the UCS for the GSR response. They were then extinguished to this stimulus complex. The procedure was the same for the two experimental groups except that 12 times during the training the light alone (or, for the other group, the tone alone) was presented without reinforcement.

The results of the experiment support the hypothesis, for the resistance to extinction of the experimental groups was significantly ($p = .05$) higher than that of the control group.

References

1. Haggard, E. A., & Gerbrands, R. An apparatus for the measurement of continuous changes in palmar skin resistance. *J. exp. Psychol.*, 1947, **37**, 92–98.
2. Kendall, M. G. *The advanced theory of mathematical statistics*. Vol. 2. London: Charles Griffin, 1940.
3. Wickens, D. D., Schroder, H. M., & Snide, J. D. Primary stimulus generalization of the GSR under two conditions. *J. exp. Psychol.*, 1954, **47**, 52–56.

(Received July 4, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

EFFECT OF VARIED REINFORCEMENT ON SPEED OF LOCOMOTION

FRANK A. LOGAN, EILEEN M. BEIER, AND ROBERT A. ELLIS

Yale University[1]

In the typical learning experiment, the aspects of reinforcement (delay, magnitude, etc.) remain constant throughout any series of rewarded trials. In the natural environment, however, the aspects of reinforcement typically vary from trial to trial. This paper reports the results of two main experiments, one in which the magnitude of reinforcement was varied while the other aspects of reinforcement were held constant, and a parallel experiment in which only the delay of reinforcement was varied.

The variation in an aspect of reinforcement may follow any of an infinite number of probability distributions. We chose to use the simplest of these: a two-point distribution with each value being equally likely. The varied magnitude group therefore received a large magnitude on a random half of the trials and a small magnitude on the remaining half. Similarly, the varied delay group was irregularly alternated between a short and a long delay of reinforcement. It will be convenient to refer to the large magnitude and the short delay as "preferred" values, by which we mean only that *S*s given a choice will typically select a large to a small magnitude and a short to a long delay.

In order to evaluate the performance of the varied reinforcement groups, two control groups for each were necessary. One control group received a constant reinforcement equal to the preferred of the two values for the varied group. The second control group received a constant reinforcement equal to the mean of the two values for the varied group.

The purpose of these experiments, then, was to observe whether groups for which reinforcement varied between two equally likely values would perform like (*a*) the mean groups which were, on the average, as well off but received their reinforcement in equal installments, or like (*b*) the preferred groups which always received the preferred of the two values for the varied groups.

No study of varied magnitude of reinforcement has been found in the literature. An experiment by Crum, Brown, and Bitterman (2) partially duplicated the varied delay of reinforcement in our design. They trained two groups of rats in a short-alley maze, one group receiving immediate reinforcement on each trial, and the second receiving immediate reinforcement on a random half of the trials and a 30-sec. delay of reinforcement on the remaining half. Because their experiment was designed for other purposes, a large portion of the learning was accomplished while both groups were receiving constant immediate reinforcement prior to the introduction of the variation in delay for the one group, which makes it difficult to conclude concerning the learning and eventual asymptote under the varied delay conditions. Further, the mean delay control group was unnecessary for their purposes, but should be available within the same design if a comparison

[1] This research was supported in part by a grant from the National Science Foundation.

between two such conditions of reinforcement is to be made.

The many experiments on partial reinforcement involve both varied delay and varied magnitude. The delay and magnitude aspects of reinforcement are not independent at nonreinforcement and the effects of varied delay and varied magnitude are therefore confounded in the partial reinforcement design. However, these studies provide possible evidence upon which to base a prediction about the present experiments. One may, on the one hand, assume that having the reward delayed a little longer on half of the trials (varied delay) is at least as good as never getting it (partial reinforcement). Similarly, one may assume that getting a somewhat smaller amount of reward on half of the trials (varied magnitude) is at least as good as getting nothing (partial reinforcement). On these assumptions, the varied reinforcement groups should perform at least as well as a partial reinforcement group would. Since the partial reinforcement training procedure typically results in asymptotic performance very slightly inferior to continuous reinforcement (4), one would predict that the varied reinforcement groups will perform almost as well as their preferred control groups and therefore better than their mean control groups.

Another approach to the prediction of performance under varied reinforcement would be to assume that *S* would run at the average of the speeds to the preferred and nonpreferred rewards. For example, when delay is varied between 1 and 9 sec., *S*'s speed would be determined by finding his speed to a constant 1-sec. delay and his speed to a constant 9-sec. delay and then taking the mean of these two speeds. If we assume that speed is a negatively accelerated increasing function of magnitude (1) and a decreasing function of delay (5), the speed of the mean of the two delays is *below* the mean of the speeds of the two delays, whereas the speed of the mean of the two magnitudes is *above* the mean of the speeds of the two magnitudes. That is, relative to the mean control groups, varied delay would do better than varied magnitude.

The present design can be subjected to a micromolar Hullian type of analysis (3) by use of the fractional anticipatory goal response (r_g). Spence (6) has suggested the use of r_g as the basis for incentive motivation (K), and the prediction requires estimating the relative amplitude of r_g conditioned to the cues immediately preceding the response under varied reinforcement.

The varied delay procedure involves the partial reinforcement conditioning of the same r_g to two different ages of the stimulus trace, i.e., a short delay trial reinforces r_g at the short age and nonreinforces it at the long age, while the reverse is true on long delay trials. Since r_g will grow almost as high under partial as under continuous reinforcement, r_g will become conditioned to both ages to almost the same degree as if reinforcement were constant at either delay. Since the same r_g is conditioned to both ages, their generalized reaction tendencies will summate behaviorally yielding a resultant r_g comparable to that produced by constant preferred delay reinforcement.

In contrast, the varied magnitude training procedure involves the partial reinforcement conditioning of two different r_g's (that appropriate to the large and that appropriate to the small magnitude) to the same age of the stimulus trace. Again the two

r_g's will grow almost as high as continuous reinforcement at either magnitude, but they are incompatible and will compete. The result will be a compromise r_g intermediate between them and comparable to that produced by constant mean magnitude reinforcement.

That is, varied delay involves conditioning the same r_g to two ages of the trace and the response tendencies summate; while the varied magnitude involves conditioning different r_g's to the same age of the trace, and they compromise. The varied delay group would therefore be expected to perform more like their preferred control group while the varied magnitude group would approximate their mean control group.

METHOD

Apparatus.[2]—The laboratory is a windowless, air-conditioned room, with a hung sound-tile ceiling. The walls are painted flat green; all furniture and equipment in the room are flat black. The rats are housed in the laboratory in individual cages: 7-w. red lights mounted above the cages are illuminated while *S*s are being run, the laboratory otherwise being in complete darkness during this time except for the lights on the apparatus itself. The home cages are separated from the rest of the laboratory by a plywood and cloth partition.

All relays, clocks, and timers (except the feeding mechanism) are isolated in Celotex boxes which are, in turn, mounted on sponge rubber inside second Celotex boxes. Any muffled sounds escaping from these boxes are effectively masked by having the air-conditioning unit in operation when the rats are being run.

The double-alley maze is suspended in a large box five sides of which are Celotex. The glass top of this box is covered with one-way vision cellophane, and is divided into three sections each provided with hinges and magnetic door catches to permit introduction and removal of *S*. The inside of this box is median gray. A sawdust pan rests 12 in. below the alleys.

The floor plan of the double-alley maze is shown in Fig. 1. The starting box and triangular choice point are median gray leading to parallel alleys one of which is flat black and the other flat white. The sides of the maze are 4 in. high; the start and runway sections are 4 in. wide. The sides are 1 in. white pine while the top and bottom are ¼-in. hardware cloth, all painted the appropriate brightness. Both sides of each alley are lined with glass.

Weighted, appropriately painted, metal guillotine doors are situated as shown in Fig. 1. The food troughs cover the entire ends of the alleys and are 1 in. high to prevent *S* from seeing the contents before breaking the last phototube beam.

The alleys are illuminated by five shielded 7-w. bulbs mounted on one side of the maze; these bulbs also serve to activate the phototubes. A 7-w. red bulb is mounted centrally above the food troughs.

There are two duplicate feeding mechanisms,[3] one for each alley. Only one is used on any single trial, depending upon the alley in which *S* runs. A single description will accordingly suffice.

A set of 18 graduated lengths of Saran tubing is mounted on a carriage. Each Saran tube is equipped with a pin to contain the pellets loaded therein. A single brass tube leads from directly below this carriage to the food cup. A reversible motor, with several external gear ratios, moves the carriage by steps in either direction; the motor starts to run when the first phototube beam is broken, and stops when the last is broken. That number of pellets loaded in that Saran tube stopping directly over the brass tube is delivered to *S*.

This food delivery is accomplished automatically when the pin in the appropriate Saran tube is removed by a bell solenoid. The bell solenoid is, in turn, activated by a Hunter Decade Interval Timer with which timing of the food delivery is easily and reliably fixed. The set of a selector switch determines which of the phototube relays activates the Hunter Timer. In this manner, reinforcement can be given at any desired time interval after *S* passes any desired point in the maze.

Reward consists of uniform pellets prepared by the P. J. Noyes Company, Lancaster, New Hampshire, primarily of Purina mash and weighing .045 gm. each. By loading the Saran tubes with varying numbers of pellets, the magnitude of reinforcement can be made to increase, or decrease, at any of a variety of rates, with *S*'s speed.[4] When the magnitude is pre-

[2] The apparatus is described in more detail than necessary for present purposes in order to avoid repetition in later reports.

[3] The writers are indebted to Gus Ogren for building the feeding mechanisms.

[4] For example, if the gears are set to move the carriage at 2 steps/sec., and if the tubes are

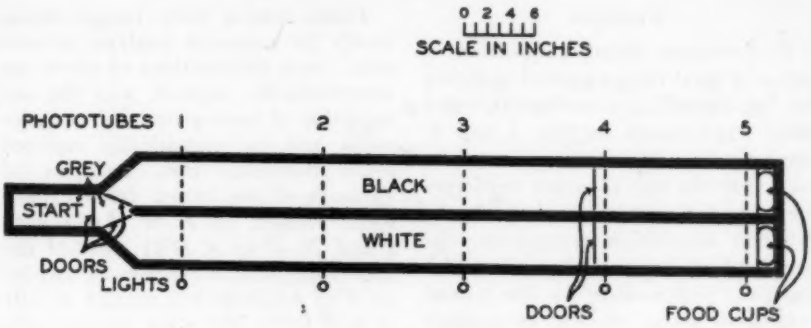

FIG. 1. Floor plan of double-alley maze.

determined independently of *S*'s speed, the motors can be set to run without moving the carriage so that only a single Saran tube need be loaded.

There are six Standard Electric Timers which measure: (*a*) time from opening the start door to first phototube beam, (*b* through *e*) running time in successive 1-ft. sections of the alley, and (*f*) running time from the first to the last phototube beam, i.e., total running time providing an independent check upon the sum of the times recorded by Clocks *b* through *e*. These clock readings are recorded and transformed into simple reciprocals to provide the response speed measure in terms of which all results are stated.

There are also available (*a*) two gray delay boxes, 18 × 18 × 4 in. in which *S*s are placed individually both before and after running the alley, (*b*) a scale for weighing *S*s, and (*c*) a balance for weighing individually the daily food ration.

Subjects.—Each experiment employed 24 black hooded rats obtained from the Sennett Antique and Pet Shop, Tiffin, Ohio. All *S*s were experimentally naive, and were between 90 and 120 days of age at the beginning of the experiment.

Training procedure.—Upon arrival, each *S* was placed in an individual living cage, and maintained on an ad libitum feeding schedule for at least four days. The *S*s were then placed on a 24-hr. feeding rhythm for at least five days before the experiment began, during which time *S*s were handled and weighed daily.

Six *S*s were run on each of four successive days for the first trial, and eight *S*s on each of the following three days for the second trial. All *S*s received one pellet of wet Purina mash in the alley on these first two trials. Thereupon all *S*s ran one trial per day according to their prescribed conditions of reinforcement in a counterbalanced running order. At least 3 min. before it was to be run, *S* was placed in one of the delay boxes available. It was then introduced into the starting box where it was detained about 20 sec. before the starting door was raised. The *S* was removed after the food reward was consumed (or after about 15 sec. on nonreward trials) and placed in a second delay box for at least 3 min. before being returned to his home cage. The day's ration, balanced for magnitude received in the alley, was given at least 6 min. later.

Experimental design.—There were three groups of eight *S*s in each experiment.[8] (A) Varied magnitude experiment (all *S*s received a constant 4-sec. delay of reinforcement): (1) 9-pellet magnitude (preferred magnitude group); (2) 5-pellet magnitude (mean magnitude group); (3) 9-pellet magnitude on a random half of each block of ten trials and a 1-pellet magnitude on the remaining half (varied magnitude group). (B) Varied delay experiment (all *S*s received a constant 4-pellet magnitude of reinforcement): (1) 1-sec. delay (preferred delay group); (2) 5-sec. delay (mean delay group); (3) 1-sec. delay on a random half of each block of ten trials and a 9-sec. delay on the remaining half (varied delay group).

loaded consecutively with 2, 4, 6, etc. pellets, then a response time of at least 1 but less than 1½ sec. receives 4 pellets, a response time of at least 1½ but less than 2 sec. receives 6 pellets, etc. By varying the speed and direction in which the carriage moves and/or by varying the number of pellets in each Saran tube, a large variety of magnitude or reinforcement conditions of reinforcement is available.

[8] One rat was lost from the varied delay group before acquisition was complete.

Results

Performance during acquisition in terms of total running speed is shown for the varied magnitude and varied delay experiments in Fig. 2 and 3, respectively. In both cases it will be seen that the two constant reinforcement control groups reached clearly different acquisition asymptotes, the preferred groups exceeding the mean groups. Performance by the varied magnitude group reached an acquisition asymptote below that of its preferred control group and comparable to that of its mean control group. In sharp contrast, however, performance by the varied delay group improved at about half the rate of its preferred control group but eventually reached an asymptote above its mean control group and comparable to its preferred control group.

These results were tested statistically by means of analysis of variance, since distributions of speed are approximately normal, and the assumption of homogeneity of variance could not be statistically rejected using Bartlett's test. Analysis of variance of the varied delay experiment yielded an F of 11.623 with 2 and 20 df ($p < .001$); that of the varied magnitude experiment yielded an F of 4.419 with 2 and 21 df ($.01 < p < .05$). We place greater confidence in the results of the varied magnitude experiment than suggested by this p value because it was the second of this design which was run. The first used a one- and seven-pellet varied magnitude at a 2-sec. delay, and while the results failed to reach conventional levels of statistical significance, the varied magnitude group again closely duplicated the mean magnitude control group, and both of

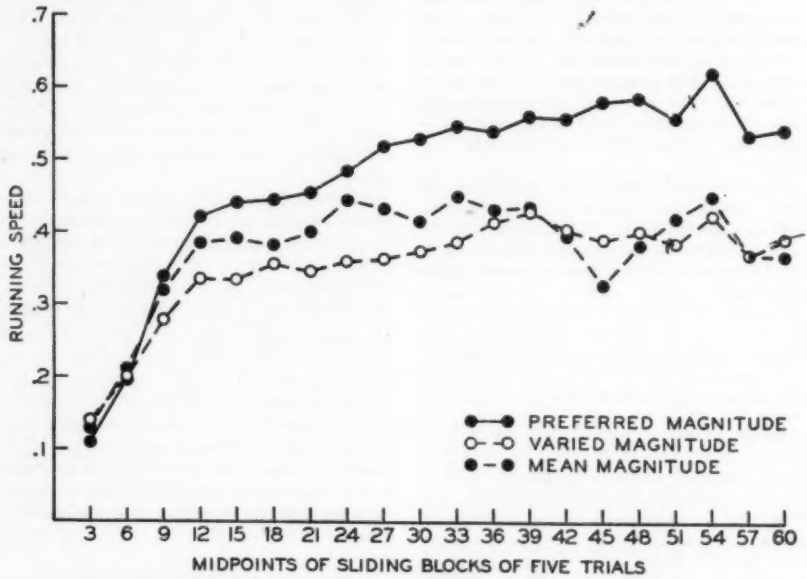

Fig. 2. Total running speeds, varied magnitude experiment.

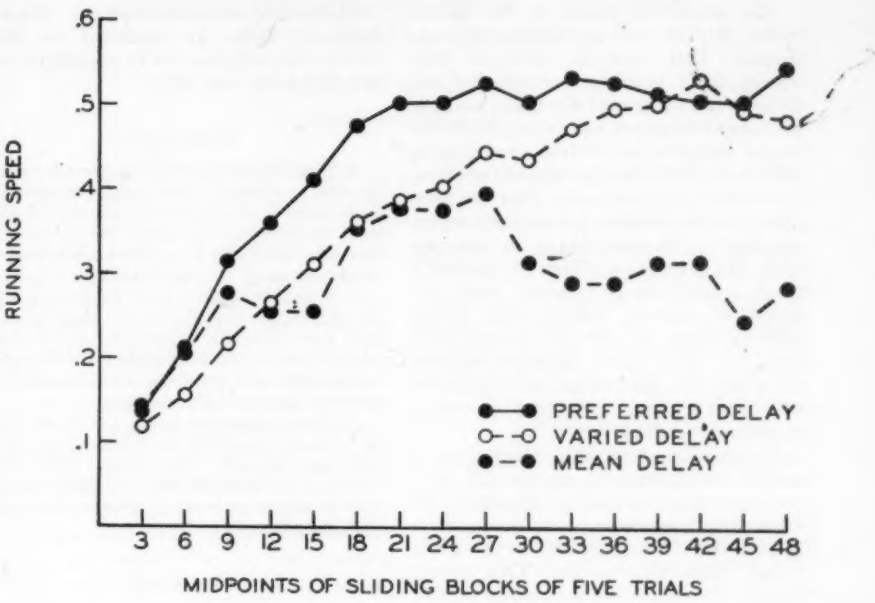

FIG. 3. Total running speeds, varied delay experiment.

these were well below the preferred magnitude control group.

Subsequent t tests show that the preferred delay control group was superior to the mean delay control group ($t = 4.676$, $p < .001$); that the preferred magnitude control group was superior to the mean magnitude control group ($t = 2.756$, $.01 < p < .02$); that the varied delay group was superior to its mean delay control group ($t = 3.345$, $p < .01$); and that the varied magnitude group was inferior to its preferred magnitude control group ($t = 2.344$, $.02 < p < .05$).

Comparable results appeared in the starting speed data; separate graphs of these are therefore not shown.

DISCUSSION

The prediction based upon the results of partial reinforcement studies failed to find confirmation in that the varied magnitude group performed significantly below its preferred control group.

The prediction based upon the assumption that performance under varied reinforcement would be at the average of the speeds to the preferred and non-preferred rewards is qualitatively confirmed by our findings. However, it deviates quantitatively in that both varied groups would not be expected to perform as well as they did, i.e., the varied delay prediction falls well below the preferred control whereas we found these two comparable, and the varied magnitude prediction falls below the mean control whereas we found these two comparable. Of course, this approach could be made to conform more closely quantitatively by use of some average more heavily weighting the fast speeds (e.g., geometric mean). Indeed, a "psychological average" might result from systematic exploration of a number of varied reinforcement conditions.

The prediction made by the micromolar Hullian type of analysis was confirmed. This analysis assumed that varied delay involves conditioning the same r_g to two ages of the trace, and the response tendencies summate; while the varied magnitude involves conditioning different r_g's to the same age of the trace, and they compromise. This is contrasted with partial reinforcement where no small r_g is conditioned to compete with the large one. That is, giving a small magnitude is actually worse, in terms of running speed, than giving nothing, since it provides a competing small r_g. This theory therefore delivers the prediction that partial reinforcement will result in faster running than varied magnitude reinforcement.

In order to test this prediction, a partial reinforcement group ($N = 7$) was run and was treated identically to the varied magnitude group except that *S*s received nothing on half of the trials rather than one pellet. This group reached an acquisition asymptote of .641 which, as predicted, was above the .405 asymptote of the varied magnitude group. In order to be sure that this result did not occur simply because the partial reinforcement group was by chance a fast one, *S*s were switched to varied magnitude reinforcement after asymptotic performance had been attained under partial reinforcement. That is, one pellet was now given on half of the trials where previously there had been nothing. According to the theory, this should provide opportunity for the conditioning of a small r_g which, as it is acquired, should compete with the large r_g already conditioned, and hence produce a gradual response decrement.

Only 40 trials (20 small reward trials) were run after the switch to varied magnitude, and the curve had not, by then, reached a clear asymptote. It had, however, fallen as predicted to .586 which was sufficient to be significant at the .025 level (one tail).

Summary

1. When the delay of reinforcement is varied between two equally likely values, asymptotic performance is comparable to constant reinforcement at the preferred of the two values and significantly superior to constant reinforcement at the mean of the two values.

2. When the magnitude of reinforcement is varied between two equally likely values, asymptotic performance is comparable to constant reinforcement at the mean of the two values and significantly inferior to constant reinforcement at the preferred of the two values.

3. These contrasting results can be handled by a theory employing the fractional anticipatory goal response. This theory yields the confirmed prediction that partial reinforcement leads to higher performance than varied magnitude reinforcement.

References

1. Crespi, L. P. Amount of reinforcement and level of performance. *Psychol. Rev.*, 1944, **51**, 341–357.
2. Crum, J., Brown, W. L., & Bitterman, M. E. The effect of partial and delayed reinforcement on resistance to extinction. *Amer. J. Psychol.*, 1951, **64**, 228–237.
3. Hull, C. L. *A behavior system.* New Haven: Yale Univ. Press, 1952.
4. Jenkins, W. L., & Stanley, J. C. Partial reinforcement: a review and critique. *Psychol. Bull.*, 1950, **47**, 193–234.
5. Perin, C. T. A quantitative investigation of the delay-of-reinforcement gradient. *J. exp. Psychol.*, 1943, **32**, 38–51.
6. Spence, K. W. Theoretical interpretations of learning. In C. P. Stone (Ed.), *Comparative psychology.* New York: Prentice-Hall, 1951. Pp. 239–291.

(Received for early publication September 23, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

GENERALIZATION OF SERIAL POSITION IN ROTE SERIAL LEARNING[1]

RUDOLPH W. SCHULZ[2]

Northwestern University

Underwood (3) has suggested that the serial position factor might be an important source of generalization in serial learning. According to his reasoning, if there is generalization among serial positions, items at the beginning and end of a list could be expected to be more discriminable by virtue of the greater discreteness of these positions. The center items are, contrariwise, less discriminable, assuming the positions they occupy have greater homogeneity (i.e., *S* is more likely to be confused whether an item is in the sixth or seventh position than he would be in designating whether the same item belonged in the first or second position in a ten-item list). In short, the less discrete middle positions should generalize more than the relatively discrete end positions. Thus, the following prediction is made from the generalization hypothesis: Differential generalization among the respective serial positions should affect *S*'s ability to designate the ordinal number of the positions occupied by the items in a serial list, after having learned that list. Furthermore, positions in the center of the list should be more difficult to designate than positions at the beginning and end of the list.

[1] I gratefully acknowledge the generous assistance and counsel given me by Dr. Benton J. Underwood and Dr. Carl P. Duncan in carrying out this research and in the preparation of the manuscript for publication. Dr. Quinn McNemar's advice concerning the statistical treatment of the data is sincerely appreciated. I also wish to thank Mr. Jack Richardson for his many helpful suggestions.

[2] Now at Stanford University.

The present experiment is an attempt to test the hypothesis of generalization of serial position. For our purposes, generalization of serial position is operationally defined as "position confusion" in terms of ordinal number (e.g., *S* designating an item actually in Position 5 as being in Position 3).

PROCEDURE

General.—The *S* was read standard anticipation learning instructions. Then, a ten-item serial nonsense syllable list was presented on a Hull-type memory drum at a 2-sec. rate with an 8-sec. intertrial interval. Immediately after *S* reached the criterion of list mastery specified for his condition, the drum was stopped and *S* was read the following instructions:

"I am now going to call aloud, one at a time in random order, each of the ten syllables from the list you have just learned. Immediately after I call off an item I want you to give me the number of the position in the list which this syllable occupied, the position of the first syllable being Number 1 and that of the last syllable Number 10. You must answer very quickly since I will call the items out at a fairly rapid pace. Further, you must give me a number for every item I call out, if you do not remember an item just do the best you can; however, you must always give me a number immediately after I have read off a syllable. Are there any questions?"

The speed set in these instructions was necessary to prevent *S* from "re-reciting" the list in designating an item's position. The *E* also practiced at length before beginning the experiment so that the time lapse between *S*'s response and *E*'s presentation of the next syllable was negligible. Hence, this position-naming response should be a fairly accurate estimate of *S*'s knowledge of item position relatively unconfounded by his level of list mastery per se. The position-naming response then constitutes the primary dependent variable throughout this study. The procedure thus far described was the same for all conditions.

Specific conditions.—The five conditions of

TABLE 1

MEAN TRIALS TO LEARN AS A FUNCTION OF INTRALIST SIMILARITY AND LEVEL OF LIST MASTERY

Cond.	*N*	Intralist Similarity	Criterion of Mastery	Mean	σ_m
I	30	Low	5/10 correct	8.60	0.80
II	30	Low	10/10+5 trials	27.63	1.74
III	30	Low	10/10+20 trials	52.20	3.17
IV	30	High	5/10 correct	11.73	1.43
V	30	High	10/10+5 trials	36.87	2.28

the experiment are summarized in Table 1. A list of low intralist similarity was used in Cond. I, II, and III. Thus, any effect due to the serial position factor should be shown more clearly, since generalization along other dimensions is held to a minimum in these conditions.

It was, of course, impossible to remove all the formal similarity among the syllables composing the low similarity list. There was the unavoidable replication of vowels, the syllables were made from letters of the alphabet, etc. Differences in position naming at the respective serial positions could result from this residue of formal similarity in Cond. I, II, and III. However, if differences in position naming noted with the low similarity list fail to change significantly when learning is under identical conditions with a list of high intralist similarity, as in Cond. IV and V, it is unlikely that these differences were produced by the formal characteristics of the list per se.

Degree of learning was varied as indicated in Table 1. The 1- perfect-plus-5-trial criterion in Cond. II and V was adopted so that *S* would have very thorough mastery of the list before position naming began, thus hoping to minimize further the possible confounding of list knowledge and position knowledge.

Items in the middle of the list never receive as many reinforcements (correct anticipations) as the end items in serial learning. The 1-perfect-plus-20-trial criterion of Cond. III was selected so that a comparison of Cond. III with Cond. II would afford an evaluation of possible differences in position naming resulting from this inequality of reinforcement.

Originally, it was thought Cond. II, III, and V would suffice for the purposes of this experiment. However, after these conditions had been run an analysis of the data suggested the need for Cond. I and IV. Cond. V may not have provided an adequate evaluation of the effect of intralist similarity on position naming. There may not have been sufficient time for the generalizing properties of these similar syllables to recover from the high degree of differentiation established in this condition. Therefore, learning in Cond. I and IV was to a criterion of 5 of 10 correct anticipations. Differentiation was incomplete and a better evaluation of the effect of intralist similarity provided. In addition, knowledge concerning the relationship between level of list mastery and position-naming ability was also derived.

List construction.—A basic list of low intralist similarity was constructed from the following ten syllables: GAC, WEP, SOZ, BIH, XEL, KIV, DUF, TAJ, YOR, and QUN. Five vowels are used twice and all consonants except M are used once. Every syllable begins and ends with a different consonant.

A basic list of high intralist similarity was constructed from the following ten syllables: WAH, WEH, WIH, NOY, NEY, NUY, KOB, KIB, KUB, and KAB. Five vowels are used twice. Two different consonants start each of three syllables. One additional consonant starts four syllables. Two different consonants end each of three syllables. One additional consonant ends four syllables. The letter which starts a syllable always occurs with the same ending letter in all other syllables it starts.

The syllables in these two basic lists were then arranged into ten different serial orders by the method of systematic randomization (**3**, pp. 130–131) to make ten low and ten high similarity lists. In this way biasing of a position in terms of differential syllable difficulty was controlled for. All syllables have association values of 46.67% to 53.33% according to Glaze (**1**). The same letter never occurred consecutively in any of the serial orders. The number sign (#) was used as a starting cue in all lists. Only three *Ss* learned the same order (or list) in a given condition. In all, 150 undergraduate *Ss* naive in verbal learning experiments served in the experiment.

RESULTS

General.—First, it is necessary to show that the variations in similarity

and degree of learning did, in fact, produce the differences in learning required for the analysis of position naming. The mean number of trials to learn under the five respective conditions is presented in Table 1. The three criteria of list mastery selected obviously produced the expected differences in learning. A statistical test of significance was not performed.

Similarity was likewise an effective variable. The high similarity list of Cond. V required 9.24 additional mean trials to be learned to the same criterion as the low similarity list of Cond. II. This difference in mean trials to learn under Cond. II and V is statistically significant beyond the 1% level ($t = 3.23$: this and all subsequent t's evaluated with 58 df). The 3.13 mean trial difference between Cond. I and IV just fails to meet the requirements of statistical significance ($t = 1.92$). However, since the difference in trials to learn the same lists under Cond. II and V was significant, it is probably safe to assume that this difference between Cond. I and IV is also a true one which would become significant with a higher criterion of list mastery. Hence, rates of learning under Cond. I and IV, as well as those under Cond. II and V, will be considered significantly different.

The *Ss* in each condition were divided into the 15 fastest and 15 slowest learners to permit a comparison of the position-naming ability of fast and slow learners. There was essentially no difference in position naming as a function of speed of learning in any of the five conditions.

The familiar bowed curves representing difficulty in learning at each serial position for all five conditions are shown in the upper half of Fig. 1. High intralist similarity appears to steepen the bow somewhat. This steepening is most noticeable late in learning, as is shown by comparison of Cond. II and V.

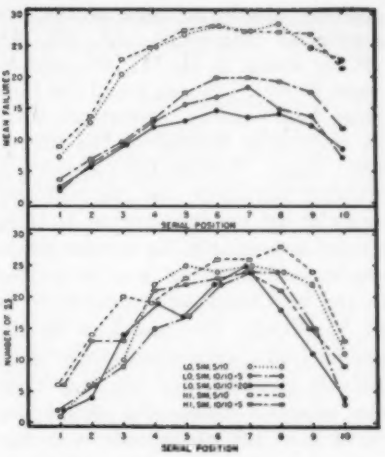

FIG. 1. Upper half: Serial position curves based on failures at each position during learning divided by mean trials to learn. Lower half: Difficulty in position naming as a function of serial position. Each point on the graph represents number of *Ss* incorrectly naming the position on the abscissa.

If difficulty in position naming is measured by the number of *Ss* failing to name the syllable in Position 1 as being in Position 1, and an item in Position 2 as being in Position 2, etc., for each of the ten respective serial positions, the curves in the lower half of Fig. 1 are derived. It is immediately seen that the position-naming curves are also "bowed curves." Difficulty in position naming likewise varies as a function of ordinal position, with the positions to the right of center being most difficult to designate.

Using the rank difference method, position difficulty as measured by mean failures in learning was shown to correlate highly with position difficulty determined by failures at

position naming in each of the five conditions (rho = .88, .93, .85, .85, .88 for Cond. I, II, III, IV, and V, respectively). Figure 1 and the foregoing high positive correlations make an important finding of the experiment clear. Neither the variation in intralist similarity nor the differing levels of list mastery appear to have altered appreciably *S*'s *relative* accuracy in position naming with respect to any particular serial position. In all conditions the positions of the middle items were consistently most difficult to name, and the end items the easiest. Thus, the dependent variable of the experiment is in line with the prediction from the generalization hypothesis.

Again comparing the position-naming curves to the bowed curves, one notes that the over-all effect produced by intralist similarity and amount of practice on the bowed curves is not apparent in the position-naming curves. Thus, the next consideration in analysis will be a more precise evaluation of position naming with regard to these variables.

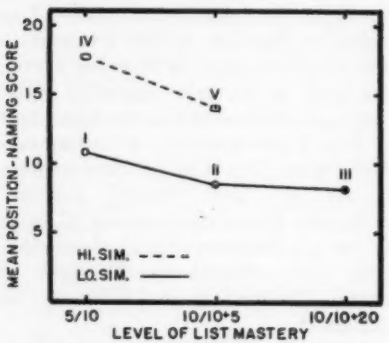

FIG. 2. Accuracy in position naming as a function of intralist similarity and level of list mastery. The roman numeral beside each point indicates the particular condition of the experiment.

Effect of intralist similarity on position naming.—A "score" representing *S*'s ability to designate position under the various conditions of the experiment was derived in terms of the absolute difference of the position *S* named from the item's actual position in the list. For example, an item actually in Position 3 which *S* named as being in Position 5 produced an absolute difference of two. The sum of the ten differences obtained in this fashion constitutes *S*'s position-naming score. Hence, a low score indicates accurate position naming and a high score inaccurate position designation. The mean position-naming scores for the 30 *S*s in each of the five respective conditions are presented graphically in Fig. 2.

Comparison of the means for the high and low similarity conditions with comparable levels of list mastery in Fig. 2 reveals that the high intralist similarity list of Cond. IV and V reduced accuracy in position naming rather substantially. The mean differences of 6.87 between Cond. I and IV, and 5.67 between Cond. II and V, are both statistically significant beyond the 1% level ($t = 5.45$ and $t = 3.69$, respectively).

In summary it is concluded that intralist similarity does effect over-all accuracy in position naming, however, as noted in connection with Fig. 1 *S*'s relative position-naming accuracy with regard to a specific ordinal position must be a function of factors other than formal item similarity. Apparently high intralist similarity simply made each position proportionately less discrete, while the relative discreteness of the end versus middle items remained fairly constant.

Effect of level of list mastery on position naming.—It is seen from Fig. 2 and Table 1 that the 24.6

additional trials received in Cond. III failed to improve position naming, when Cond. III is compared to Cond. II. Apparently *S*'s knowledge of position reaches an asymptote before the 1-perfect-plus-5-trial criterion used in Cond. II, insofar as the difference in degree of learning between Cond. II and III is an adequate sample of the dimension. The middle items in Cond. III received more reinforcements than the end items in Cond. II. Positions 6, 7, and 8 in Cond. III were reinforced 797, 853, and 826 times, respectively, while Positions 1, 2, and 3 in Cond. II were anticipated correctly 780, 659, and 572 times, respectively. Thus, the absolute number of reinforcements, defined as correct anticipations, likewise seem to contribute little to *S*'s position-naming ability.

Turning to the lower end of the degree of learning dimension and Cond. I and IV, inspection of Fig. 2 indicates that *S* has already achieved a fairly high level of accuracy in position naming, in spite of the low level of list mastery in these conditions. If Cond. I and IV are compared to Cond. II and V in Fig. 2, it is seen that position naming does improve somewhat as a function of increased list mastery. The 3.60 mean difference in position-naming score favoring Cond. V over Cond. IV is significant at the 2% level ($t = 2.34$). Likewise, the 2.40 mean difference between Cond. I and II reaches statistical significance at the 5% level ($t = 2.03$). However, it should be noted that the degree of list mastery in Cond. II and V was more than 50% higher than that achieved in Cond. I and IV; thus, the magnitude of improvement in position naming found is hardly in keeping with the amount one might have expected to find. Unfortunately, this finding is not explicable from the data of the present study.

Discussion

Accepting "position confusion" in terms of ordinal number as an operational definition of generalization of serial position, the results of the experiment are consonant with the prediction that there is differential generalization along the serial position dimension. It was shown quite conclusively that *S*'s ability to designate the ordinal number of an item's position in a serial list depended largely upon the particular position that item occupied during learning, with items in the center of the list being most difficult to designate. Empirically, the results are clear.

However, in the light of some of the findings it becomes necessary to re-evaluate the assumption that ability to name the ordinal number of an item's position is, in fact, a valid measure of generalization. The fact that increasing the degree of learning from 5 out of 10 correct to 1 perfect plus 20 trials produced only minimal changes in position-naming ability raises some questions. At least two interpretations of this finding are immediately apparent. One is that knowledge of position is learned with extreme rapidity, somewhat independent of the "ordinary" associative learning of the serial item sequence. Hovland's (2) finding that the relative amount of generalization increases rather rapidly to a maximum and then declines only slightly with further practice is favorable to such an interpretation. However, it must be remembered that Hovland's results were derived from a conditioning situation in which a minimum of "actual" learning was necessary to produce the generalized response. A second interpretation, which seems even more plausible at this point, is that *S* simply does not learn serial position as represented by ordinal number. It is hard to conceive of *S* knowing the ordinal position of items he has only responded to an average of 9 times (see

Table 1) almost as well as the positions of items he has responded to 52 times.

On the other hand, the finding that the position-naming curves of Fig. 1 are bow shaped, rather than symmetrical or some other shape, coupled with the fact that difficulty in position naming does correlate highly with difficulty in learning at the respective serial positions would seem to indicate that there is some relationship between the phenomena of serial learning and position naming. It is, however, impossible to determine the nature of this relationship from the data of the present study.

In short, the results of the experiment are inconclusive for evaluating the generalization hypothesis because of the contradictory nature of the interpretations made possible by certain findings. Hence, the response measure, position naming, may be unsuitable for a test of the hypothesis of generalization of serial position. An adequate evaluation of this hypothesis may require the development of a more direct method for measuring the possible existence of generalization along this spatial-temporal dimension.

Summary

To investigate the notion, proposed by Underwood, that there is differential generalization among serial positions, 150 *S*s were asked to learn a ten-item serial nonsense syllable list under five different conditions. Level of list mastery was varied in three ways; intralist similarity in two. Immediately after learning *S* was asked to designate the ordinal number of the serial position occupied by each syllable in the list he had just learned. Inaccuracy in naming these positions provided the operational definition of generalization of serial position. Empirically, the following was established:

1. Accuracy in position naming varied as a function of serial position. The function was bow shaped, with a particular position's relative difficulty remaining fairly constant even under widely varied conditions of learning.

2. Knowledge of serial position per se reached its maximum level of perfection somewhere between 5 out of 10 correct and 1 perfect plus 5 trials in terms of list mastery criterion.

3. Formal intralist similarity, though affecting position naming, cannot be held entirely accountable for the results of the experiment.

These findings support the generalization hypothesis in terms of the definition set forth. However, certain other findings made this definition somewhat tenuous, so it was necessary to conclude that the experiment did not adequately evaluate the hypothesis of generalization of serial position.

References

1. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.
2. Hovland, C. L. The generalization of conditioned responses. IV. The effects of varying amounts of reinforcement upon degree of generalization of conditioned responses. *J. exp. Psychol.*, 1937, **21**, 261–276.
3. Underwood, B. J. *Experimental psychology: an introduction.* New York: Appleton-Century-Crofts, 1949.

(Received July 1, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

THE SHAPE OF THE DISCRIMINATION GRADIENT FOR TWO INTRACONTINUUM STIMULUS SEPARATIONS[1]

GEORGE E. PASSEY AND PAUL N. HERMAN

University of Alabama

Two implications of Hull's (**1, 2**) theory of discrimination learning concern the shape and slope of the discrimination gradient. In his Theorem 13, Hull holds as follows:

> Theorem 13c—As the range of $\frac{1}{2}d$ is decreased, the discrimination gradient connecting S_1 and S_2 becomes less concave upward, passing into a slightly convex upward form after $d = 30$ (**2**, p. 74).
>
> Theorem 13d—As the range of d is decreased, the more nearly vertical will be a straight line drawn from the reaction potentials of S_1 and S_2 (**2**, p. 74).
>
> Where S_1 = The positive stimulus
> S_2 = The negative stimulus
> d = The intracontinuum difference between S_1 and S_2 in just noticeable differences.
> Reaction potential = A theoretical construct representing response strength.

The literature contains no evidence bearing directly upon these implicates. The present investigation attempts to secure data that allow an examination of these theorems. Following acquisition of a visual intensity discrimination, presentation of stimuli at intensities intervening between the positive and negative stimuli should yield data that will reveal the shape of the discrimination gradient. The present experiment utilizes two intensity discriminations with separations of positive and negative stimuli presumed to be less than 30 j.n.d.'s.

Theorem 13c predicts for the two discrimination gradients of the present experiment a slightly convex upward form. The greater convexity should be demonstrated for the short intracontinuum stimulus separation. For Theorem 13d a more nearly vertical line will be predicted for the short intracontinuum stimulus separation than for the long separation when the terminal reaction potentials for each gradient are connected with a straight line.

[1] This research was supported in full by the Research Committee of the University of Alabama.

Apparatus and Procedure

Albino and pigmented rats, selected at random from a colony maintained by the University of Alabama, were trained in visual intensity discrimination. An apparatus of floor plan and size comparable, except where noted, to that employed by Wilcoxon, Hays, and Hull (**6**), and described elsewhere, was utilized. The *S*s were required to proceed from a starting box through an 11-in. runway illuminated at the appropriate stimulus intensity. At the end of the runway a door was encountered which, when raided by *S*, gave access to the goal box. All *S*s were run after 22-hr. food deprivation. The apparatus was so arranged that depression of a switch by *E* activated a solenoid mechanism that simultaneously opened the door of the starting box and activated a timing device. Raising by *S* of the door separating the runway from the goal box served to interrupt a photoelectric relay which stopped the timing mechanism. Elapsed time was recorded to a precision of .05 sec. The training consisted of four phases.

Preliminary training.—A familiarization period of six days was provided for all *S*s. Each *S* was confined in the starting box for a period of 30 sec. The door was then opened and *S* given access to the illuminated runway. Half of the *S*s were exposed to a runway illumination of 64 foot-candles while the remaining *S*s were exposed to an illumination of 128 foot-candles. The intensity of illumination was determined by placing an exposure meter, especially designed to allow incident light determinations, upon the floor of the runway. On Day 1 the door separating the goal box from the runway was secured in the fully open position. On succeeding days it was lowered approximately one-fifth of the distance per day so that on Day 6 *S* was required to open the door from the completely closed position

to gain access to the goal box. On each of the six days a container of moistened Purina Lab Chow was placed in the goal box. A maximum of 5 min. was allowed in the runway; if *S* failed to enter the goal box at the expiration of 5 min. such entry was forced by *E*. After entry to the goal box *S* was allowed access to the food reward for a period of 5 min. On Day 6, *S*s failing to enter the goal box within the 5-min. period were eliminated from further training.

Acquisition of response to positive stimulus.—Each *S* was confined in the starting box for a period of 30 sec.; the door was then opened and *S* was allowed to proceed down the runway to the goal box in the presence of the positive stimulus. If *S* failed to enter the goal box in 90 sec., the trial was terminated and *S* was removed from the apparatus. The trial was scored as requiring infinite time. If *S* entered the goal box within the time limit, the elapsed time was noted and constituted *S*'s score for the trial. Time in the goal box during this phase was limited to 50 sec. and the reward used was a moistened pellet of Purina Lab Chow weighing approximately .4 gm. Five trials were given *S* on each of eight successive days—a total of 40 trials. Trials were separated by approximately 5 min. An *S* failing to enter the goal box on ten successive trials was eliminated from further training.

Discrimination training.—For both the 128 and 64 ft.-candle groups a negative stimulus of 4 ft.-candles was employed. These distances were selected so as to provide a stimulus difference value, *d*, which would be less than 30 j.n.d.'s. In one situation Lashley (3) found a discrimination of 15% of a standard brightness. The more usual finding of both Lashley and Slater and Munn (5) places the differential threshold in the neighborhood of 50% of the standard. In terms of these findings our stimulus separations would appear to fall well short of a *d* value of 30 noted in Theorem 13c.

Ten trials were given each day for a maximum of 32 days. Of the ten daily trials, three involved presentation of the positive stimulus and subsequent reward while seven utilized the negative stimulus and nonreward, although *S* was always allowed to remain in the goal box for 50 sec. after entering it. Order of stimulus presentation was randomized for each set of ten trials. In this phase of learning an *S* who failed to enter the goal box in 60 sec. was scored as requiring infinite time and removed from the apparatus. Running times were converted into rates through the use of reciprocals and means for positive and negative stimulus conditions computed. Trials terminated by *E* and scored as infinite were assigned rates of zero. Whenever the daily mean rate for the positive stimulus trials exceeded the rate for negative trials, the daily performance was considered successful. The criterion of learning employed required that *S* perform successfully for 11 consecutive days within a 32-day maximum. Upon reaching the criterion *S* was advanced to the fourth phase.

Determinations of the discrimination gradient.—A population of 36 *S*s that reached the criterion prescribed in the previous phase was used in the gradient determinations. Groups of nine albino and nine pigmented *S*s were employed for each stimulus separation. The test series utilized a period of four days immediately following the criterion day of the previous phase. On each of the four days a total of ten trials was given. Three of the ten daily trials involved presentation of the positive stimulus and were rewarded. The remaining seven daily trials were devoted to a determination of response strength to three intervening stimulus intensities and the negative stimulus, each unrewarded. Counterbalanced orders were employed so that stimuli other than the positive were presented seven times each over the 4-day test period. The positive stimulus was presented 12 times. For the group having as a positive stimulus 128 ft.-candles, intermediate intensities of 64, 32, and 16 ft.-candles were used. The 64 ft.-candle group was stimulated with intermediate intensities of 32, 16, and 8 ft.-candles.

Results and Discussion

The data obtained in the gradient test series are summarized in Fig. 1. The median reaction latency ($_{st}R$) for the group of *S*s tested at each stimulus intensity was transformed into reaction potential ($_{s}E_{R}$) utilizing the equation of Hull: $_{s}E_{R} = 2.845\ (_{st}R)^{-.483}$ (2, p. 13), and plotted on the ordinate. In order to produce a scale for the abscissa conforming to a scale of just noticeable differences, as required by the theorem, the logarithms of the stimulus intensities were obtained. The differences between these logarithms were then plotted, taking the point of origin as the positive stimulus intensity. When the variances of the reaction potentials for their respective stimulus points were calculated, they were found to demonstrate nonhomogeneity. It was not possible, therefore,

to evaluate the departure of the gradients from linearity. Inspection of the curves reveals them to be of a convex upward form. This finding agrees with the prediction of Hull for stimulus separations less than 30. We interpret this finding as supporting Theorem 13c.

Theorem 13d holds that a straight line drawn connecting the terminal reaction potentials of each of the discrimination gradients will yield a line of greater slope for the 64–4 ft.-candle group when compared with that for the 128–4 ft.-candle group. The slope of the line for the 64–4 ft.-candle group is −.88 and that for the 128–4 ft.-candle group is −.85. While the difference in slope is in agreement with the prediction of the theorem, it is an extremely small difference. Since it is not possible to evaluate the significance of this finding, it can be considered only suggestive regarding conformity with the theorem.

Hull's Theorem 13e predicts that the net discriminatory reaction potential for the 64–4 ft.-candle group will fall below that for the 128–4 ft.-candle group at the positive stimulus point. Inspection of Fig. 1 indicates an inversion of these points. The difference between them is, however, not significant ($t = 1.33$, $df = 34$). For those stimulus separations employed we interpret our finding as contradicting Theorem 13e. It may well be, however, that the separations employed in the present experiment were not sufficiently different to provide an adequate test of this theorem. The finding of the extremely small difference in our test of Theorem 13d may also be a function of the stimulus separations chosen. Further evaluation of these theorems is necessary.

Data summarizing the acquisition of the simple running response for those 36 *Ss* reaching criterion performance are presented in Table 1. Inspection of these data indicates satisfactorily comparable performance of albino and pigmented *Ss* and absence of a stimulus intensity difference in acquisition of this response.

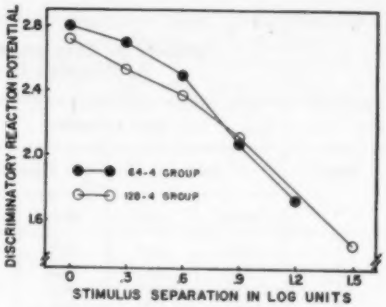

FIG. 1. Brightness discrimination gradients plotted in reaction potential at points along the stimulus continuum differing in log value as indicated.

A comparison of pigmented and albino *Ss* with regard to acquisition of the required brightness discriminations is also possible from the data obtained. The number of days necessary to reach the learning criterion was employed as an index of performance. For the 64–4 ft.-candle group the pigmented *Ss* required a mean of 19.33 days and the albino *Ss* a mean of 23.66 days while for the 128–4 ft.-candle group means of 16.77 and 21.11 days, respectively, were obtained. These data do not include those *Ss* failing to reach criterion performance. In the group having a positive stimulus of 64 ft.-candles, and hence the short stimulus separation, two albino and three pigmented *Ss* failed to reach the criterion. In the group having a positive stimulus of 128 ft.-candles, and hence the long stimulus separation, one albino *S* failed to reach criterion performance. These

TABLE 1

SUMMARY OF LEARNING DATA UTILIZING RECIPROCALS OF REACTION LATENCIES AS AN INDEX

Days	128 Ft.-Candles				64 Ft.-Candles			
	Albino		Pigmented		Albino		Pigmented	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
1	.088	.067	.063	.056	.098	.075	.149	.105
2	.288	.136	.160	.110	.252	.166	.287	.210
3	.459	.267	.406	.232	.377	.201	.400	.219
4	.552	.223	.501	.211	.564	.181	.559	.193
5	.628	.358	.550	.203	.602	.124	.715	.583
6	.545	.212	.723	.184	.662	.136	.764	.185
7	.760	.146	.809	.258	.735	.110	.715	.117
8	.734	.135	.762	.145	.741	.098	.776	.141

data were analyzed utilizing the technique of analysis of variance. Selection of the samples compared through the enforcement of the learning criterion will underestimate the true difference between the albino and pigmented *S*s.

The interaction of strains and stimuli was not significant ($F = 2.01$, 1 and 32 *df*). The F for the rate of acquisition as a function of stimulus intensity was 2.21 ($p > .05$, 1 and 32 *df*), and the hypothesis of no difference in rate of acquisition as a function of stimulus intensity is accepted. The F for difference in performance between albino and pigmented *S*s was 6.47 ($p < .05$, 1 and 32 *df*), and is considered to be evidence for a difference in the rate of acquisition of the discriminatory response for the strains employed. This result taken with the finding of Lashley (**4**) that the pigmented *S* is superior to the albino *S* in form discrimination further indicates the desirability of employing the pigmented *S* where visual tasks are to be investigated.

SUMMARY

Two groups of *S*s consisting of nine albino and nine pigmented *S*s each reached a learning criterion in acquiring a differential running response to two intensities of illumination. One group was exposed to a large intracontinuum stimulus separation, 128 versus 4 ft.-candles, while the second group was exposed to a short intracontinuum stimulus separation, 64 versus 4 ft.-candles. The criterion employed was reached by the pigmented *S*s more rapidly than by the albino *S*s. The probability of the occurrence of the difference due to chance was $<.05$.

After reaching the criterion *S*s were tested for performance on intensities lying between the positive and negative stimuli in order to ascertain the shape of the gradient of reaction potential separating them. Hull's theorem 13c, which holds that for a stimulus separation d less than 30 a convex upward discrimination gradient will be obtained, is supported by our data.

Evidence suggestive of conformity with Hull's Theorem 13d was obtained. With increasing stimulus separation a straight line drawn between the positive and negative stimulus reaction potentials is more nearly vertical for the short stimulus separation than for the larger stimulus separation. Also, data contradicting Hull's Theorem 13e were obtained. The possibility of inadequate separation of stimuli was suggested as limiting the generality of this finding and that for Theorem 13d.

REFERENCES

1. HULL, C. L. Simple qualitative discrimination learning. *Psychol. Rev.*, 1950, **57**, 303–313.
2. HULL, C. L. *A behavior system.* New Haven: Yale Univ. Press, 1952.

3. LASHLEY, K. S. The mechanism of vision: II. The influence of cerebral lesions upon the threshold of discrimination for brightness in the rat. *J. genet. Psychol.*, 1930, **37**, 461–480.
4. LASHLEY, K. S. The mechanism of vision: III. The comparative visual acuity of pigmented and albino rats. *J. genet. Psychol.*, 1930, **37**, 481–484.
5. SLATER, J. E., & MUNN, N. L. A note on brightness vision in the white rat. *J. comp. Psychol.*, 1932, **13**, 273–277.
6. WILCOXON, H. C., HAYS, R., & HULL, C. L. A preliminary determination of the functional relationship of effective reaction potential (sE_R) to the ordinal number of Vincentized extinction reactions (n). *J. exp. Psychol.*, 1950, **40**, 194–199.

(Received June 4, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

A DEVELOPMENTAL STUDY OF TRANSPOSITION

HAROLD W. STEVENSON, IRA ISCOE, AND CLAUDIA McCONNELL

The University of Texas

In several recent studies with preschool children (**1, 4**) an attempt has been made to determine the changes that take place in discrimination learning and in transposition with increasing chronological and mental age. The results indicate that older preschool *Ss* learn a size discrimination with greater ease than younger *Ss*, and that transposition to new sets of stimuli is greater for the older *Ss*.

The purpose of the present experiment is to extend the age range beyond the preschool level and to explore the changes that occur in the learning and in the transposition of a size discrimination from the preschool through the adult years.

METHOD

Subjects.—The *Ss* were obtained from (*a*) nursery schools in Austin, Texas; (*b*) public schools in Victoria, Texas; and (*c*) the University of Texas.[1] The public school *Ss* (20 each from Grades 2, 5, 8, and 10) were selected on the basis of having IQ's between 90 and 110 as determined the preceding year by a group intelligence test (**5**). In each group there was an equal number of boys and girls and an equal number of *Ss* having IQ's below and above 100. The nursery *Ss* ($N = 20$) were selected on the basis of CA alone (4 yr. 0 mo. to 4 yr. 11 mo.), since test results were not available for the determination of their MA. The college sample consisted of 20 undergraduate volunteers who were not psychology students. Ratings of MA were unavailable for these *Ss*, and no attempt was made to select *Ss* of a particular CA. The preschool and college groups were equally divided with regard to sex. No *S* had previously participated in psychological experiments.

TABLE 1

THE MEAN CA, MA, AND IQ OF THE *Ss* ACCORDING TO SCHOOL GRADE

Grade	*N*	CA (Years)		MA (Years)		IQ Mean
		Mean	*SD*	Mean	*SD*	
Preschool	20	4.6	.2			
2	20	7.8	.2	7.8	.5	100
5	20	11.0	.5	11.1	.8	101
8	20	13.8	.5	13.8	.6	100
10	20	16.2	.4	16.3	.6	101
College	20	20.0	1.8			

Materials.—The equipment consisted of a screen, a 14- × 9-in. board, and four 3-in.-sq. blocks. The screen was used to prevent *S* from observing *E* arrange the blocks. Along the midline of the board were two holes, 1½ in. in diameter and separated by 3½ in., which served to contain the goal object (an eraser). On three sides of each hole was a raised wooden guide to facilitate the proper placement of the blocks over each hole. The board and blocks were painted flat black. Mounted in the center of each block was a white cardboard square, the stimulus to be discriminated. The top of each block was covered with a piece of clear Lucite to prevent soiling through handling. The areas of the squares on Blocks 1 to 4 were ¼, 1, 2, and 4 sq. in., respectively.

Procedure.—The *S* was seated across the table from *E* and *E* demonstrated how the blocks fitted over the holes. The *S* was then told that *E* would hide the eraser and that the point of the "game" (or "experiment" for college *Ss*) was for *S* to find the eraser as often as possible.

All *Ss* were trained on Blocks 2 and 3. The position of the blocks on the board was varied randomly, but the sequence was the same for all *Ss*. For half of the *Ss* the smaller stimulus was correct and for the other half the larger stimulus was correct. The training trials were continued until nine consecutive correct responses were made. Immediately after this criterion was met the test trials were begun without comment. The *Ss* for whom the smaller stimulus had been correct on the training trials were tested on Blocks 1 and 2, and the *Ss*

[1] We acknowledge the kind cooperation of the staff at the Jeffrey Nursery School and the Jack and Jill Nursery School, and of the principals and teachers of the Victoria, Texas, public schools.

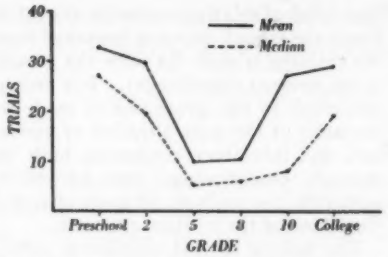

FIG. 1. Mean and median number of trials required to learn the original discrimination.

for whom the larger stimulus had been correct were tested on Blocks 3 and 4. The eraser was again under the smaller or larger block. The test trials were continued until five consecutive correct responses had been made. Following the test trials all *Ss* were asked how they knew where to find the eraser.

RESULTS AND DISCUSSION

Training trials.—The mean and median number of trials required to reach the training criterion (excluding the nine criterion trials) may be seen in Fig. 1. There is a marked curvilinear relationship between trials and developmental level. The tenth grade and college *Ss* learned the discrimination only slightly faster on the average than the preschool and second grade *Ss*, and the *Ss* in Grades 5 and 8 learned it with the greatest ease of all. When tested by Festinger's *F* test (2) both the drop in means from Grade 2 to 5 and the rise from Grade 8 to 10 are significant beyond the .001 level of confidence ($F = 2.99$, $df = 52,32$; $F = 2.59$, $df = 31,36$).

The decrease in trials through Grade 5 is in line with previous findings showing improved speed of learning with increased chronological age. The rise in trials after this is unexpected, but may be understood on the basis of *Ss*' verbalizations at the end of the experiment. The Grade 10 and college *Ss* indicated that they expected a much more difficult problem and tried to find a solution in the form of a complex sequence of positions or stimulus patterns. Or, in some cases, *Ss* reported that they assumed success depended upon chance and that there was no solution at all. The set to solve a more complicated problem or to attain success by guessing was related to the slowness of these *Ss* in learning the simpler, successful response.

Test trials.—Three different measures of transposition were available: (*a*) the mean number of relative responses on test Trials 1–5; (*b*) the proportion of *Ss* making a relative response on the initial test trial; and (*c*) the proportion of *Ss* making consistent relative test responses from the initial test trial through the test criterion. As may be seen in Fig. 2 and 3, there is a tendency for the groups who learned the initial discrimination most readily (Grades 5 and 8) to make the greatest number of relative test responses. However, chi-square tests of over-all differences between grades yield p values greater than .20 for each criterion, indicating that the incidence of relative responses does not differ significantly throughout the age range tested.

At no age level was transposition perfect for all *Ss*, even though at nearly every age level the performance was above chance. The significance of the departure from chance can be

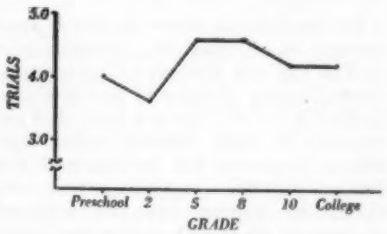

FIG. 2. Mean number of relative responses made on the first five test trials.

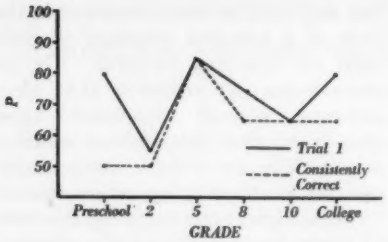

FIG. 3. Proportion of *Ss* making relative responses on the first test trial and proportion of *Ss* making consistent relative responses through the test criterion.

determined only for the responses on the initial test trial, since it is impossible, with the testing method used, to determine the chance level for the other measures. On the initial test trial the performance of all groups was above chance except for Grades 2 and 10 ($CR > 1.96$, $p < .05$).

At the end of the experiment all *Ss* were able to verbalize the relationship of stimulus size to correct solution, except for 12 *Ss* in the preschool group. These *Ss* learned the discrimination on the training trials somewhat more slowly than the verbal preschool *Ss* did (37.7 vs. 25.2 trials), but the difference is not statistically significant ($F = 1.50$, $p > .05$). On the test trials the nonverbal *Ss* showed a lower incidence of relative choices than their verbal agemates, but again none of the differences was significant.

In the light of these results it is of interest to consider the hypotheses of Koffka (3) and Kuenne (4) concerning developmental changes in transposition. Koffka (3, p. 141) has proposed that the tendency to make absolute rather than relative responses will increase with CA because of the improvement in *S*'s ability to discriminate the absolute qualities of the stimuli and to retain this discrimination on the test trials. Kuenne on the other hand has predicted that regardless of *S*'s CA a consistently high level of relative responses should be found for stimuli one step removed from the training stimuli (as were the stimuli in the present experiment). For young, preverbal *Ss* this prediction is made on the basis of the generalization of excitatory and inhibitory tendencies built up through conditioning, and for older, verbal *Ss* it is made on the basis of verbal mediation of the relative response.

The failure to find significant differences in the frequency of relative choice is contrary to the prediction of Koffka, but offers some support to that of Kuenne. However, the imperfect responses on the initial test trial and the inconsistent behavior on subsequent trials indicate that the behavior was not guided solely by an implicit verbalization of the previously correct solution. Thus, although language may aid transfer, a simple verbal mediation hypothesis is an oversimplified description of the transfer behavior of the verbally sophisticated *S*.

SUMMARY

Six groups of 20 *Ss* were selected from preschool, Grades 2, 5, 8, and 10, and college in order to explore the developmental changes in the learning of a size discrimination and in the transposition to a set of stimuli one step removed from the training pair. A decrease in learning trials was found through Grade 5, but a significant increase appeared following Grade 8. The incidence of transposition did not differ significantly among the groups.

REFERENCES

1. ALBERTS, E., & EHRENFREUND, D. Transposition in children as a function of age. *J. exp. Psychol.*, 1951, **41**, 30–38.
2. FESTINGER, L. A statistical test for means of samples from skew populations. *Psychometrika*, 1943, **8**, 205–210.
3. KOFFKA, K. *The growth of the mind.* (Trans. by R. M. Ogden.) New York: Harcourt Brace, 1925.
4. KUENNE, M. R. Experimental investigation of the relation of language to transposition behavior in young children. *J. exp. Psychol.*, 1946, **36**, 471–490.
5. SULLIVAN, E. T., CLARK, W. W., & TIEGS, E. W. *California Test of Mental Maturity.* Los Angeles: California Test Bureau, 1951.

(Received June 30, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

THE INFLUENCE OF PAST ASSOCIATIONS UPON ATTRIBUTIVE COLOR JUDGMENTS

KATHERINE E. BAKER AND IRENE MACKINTOSH

University of Nebraska

While color judgments are most frequently studied with stimulus factors as the independent variables, or as a function of adaptation conditions of the eye, it is quite likely that motivational, cognitive, and, in general, experience factors are also determinants of attributive color judgments. One convincing demonstration of such influences is found in the now well-known experiment by Bruner, Postman, and Rodrigues (2). These investigators showed that color matches for familiar forms (i.e., tomato, lemon, etc.) were in accordance with the usual color attributes of the objects depicted even when the forms were presented as silhouettes of gray paper against a blue-green background. In their experiment all the silhouettes possessed an identical orangish hue by color contrast, and *Ss* were instructed to manipulate red and yellow sectors on a rotating color wheel to achieve the best match possible. The relationship among matches for different forms was predictable on the basis of consideration of the natural colors of the objects pictured and could not be attributed to differences in actual color of stimuli. Such an influence upon color judgments must be dependent upon past commerce with the objects such that color-form associations are learned.

The question remains, however, as to the conditions and the nature of the learning which takes place and shows its residual effects in future perceptions. The types of objects represented in the study by Bruner, Postman, and Rodrigues undoubtedly were very familiar to their *Ss*, who presumably had come in contact with the objects in their natural colors a great many times. Thus, both the forms and their colors might be expected to have become clearly differentiated from other familiar objects. One might well wonder if such a protracted series of learning trials is a necessary condition for past associations to affect color judgments. Furthermore, drive reduction of a basic sort quite probably was intimately related to many experiences *Ss* had had with objects of the type used in the previous experiment. The necessity of this condition of learning in the production of perceptual effects is as yet unestablished. In addition, it needs to be recognized that familiar objects are often encountered in terms of verbal symbols. Whether or not spoken or read verbal symbols for attributes of objects can, in and of themselves, influence the perceptions of the objects is still an unanswered question.

The present experiment purports to shed some light on these questions raised by the former investigation. Following the findings of Solomon and Postman (4) that threshold recognition times for tachistoscopically presented stimuli show great changes over frequencies of prior exposure between 0 and 25, it is hypothesized here that influences upon attributive color judgments, similar to those observed by Bruner, Postman, and Rodrigues, may be brought about by relatively few color-form exposures of unique nonsense forms. While it is

extremely difficult to make unequivocal inferences concerning human motivation, an effort was made to form the desired associations without depending upon high motivation or obvious drive reduction. Further, because of the current importance placed by Vernon (5) and others upon the influences of "naming" or labeling upon perception, it is also hypothesized that a comparable effect upon attributive color judgments may be accomplished by simply associating color names or labels with neutral gray nonsense forms.

Procedure

The procedure of the present experiment was similar to that of Bruner, Postman, and Rodrigues; the major change was the use of unique nonsense forms and a pretest situation during which color associations were learned for the unfamiliar forms.

Twenty-six *S*s, all university students with a maximum of 3 hr. credit in psychology, participated in the experiment. Upon entering the experimental room, each *S* was told that he was to be an *S* in an experiment on association and that he would be shown a number of forms paired with color names. His job was to say aloud the appropriate color name upon seeing the form alone, and he was told to check himself when the printed word appeared.

During the pretest situation six different form-color name pairs were presented 14 times each in an order determined from a table of random numbers. Each form was made up of a complex combination of straight and curved lines, and an effort was made to have each form distinctively different from each other form and, at the same time, to have no form bear a resemblance to any familiar object or shape. The forms were all about 2½ in. high and 1½ in. wide and each was cut from paper and mounted on a separate 2 × 3½-in. white card over which a piece of clear acetate was fitted. Three of the six forms were chosen to test the hypothesis that color labeling alone might affect color judgments and, therefore, these forms were always presented as neutral gray (Stoelting paper No. 19). One form was paired with the word "yellow," another with the word "orange," and the third with the word "red." These words were written in large black capital letters on separate white cards. The remaining three forms were used to present color as an attribute of form. One form was cut from yellow paper, one from orange paper, and the third from red paper (Stoelting Nos. 5, 3, and 1, respectively). Each of these forms was paired with its correct color label. Because of the possibility that the shape of the colored forms might be ignored in favor of the simpler color-color label association, identical neutral gray replicas of these three forms were also used. For 7 out of the 14 exposures of each of these forms the figure itself was colored, and for the other 7 exposures the neutral gray figures were employed, keeping the same form-color label relationships.

The apparatus used to expose the pairs of cards presented to *S* a 24 × 36-in. gray painted screen in the center of which were two covered rectangular apertures, 2 × 3½ in. in size and 1½ in. apart. Following a ready signal, the shutter behind the right-hand window was raised to expose a form lighted by a 7½-w. bulb behind the screen. After a 2-sec. interval the mechanism driven by a Bodine synchronous motor automatically raised the left-hand shutter to expose the lighted card upon which the paired color label was printed. The form and its color label remained exposed together for another 2 sec. following which both shutters lowered simultaneously. After approximately 10 sec., during which necessary card changes were made, the next trial was begun. The entire 84 exposures were given without rest periods, each trial being separated from the next by the 10-sec. interval only.

At the end of the part of the session in which forms and their color labels were exposed, *S*s were led to believe that the experiment was over. They were told so and, in addition, they were asked if they had discovered "the pattern" which was implied to exist in the actually random order of exposures. After a few minutes conversation about the pseudo explanation of the experiment, each *S* was asked if he would be willing to stay a few more minutes to help test an apparatus which *E* was setting up for another staff member. All *S*s complied with the request and they were then tested for color matches of the forms employed in the previous part of the session. The use of the same forms was explained by saying that some old cards were left over, and, since the apparatus was merely being tried out, cards which were handy were being used. To assure further that *S*s had not discovered the relationship between the two parts of the session, and therefore had reacted in terms of what was presumed to be expected of them, each *S* was questioned when the experiment was in fact over. The very rare *S* who did not admit that he had been successfully duped was dropped from the study. Care was taken to solicit each *S*'s co-

operation in keeping information about the experiment from spreading to other potential *Ss*.

For the color-matching part of the experiment, *S* was seated in front of a gray painted screen containing an 8-in. diameter aperture backed by a sliding door. He was instructed that he would be asked to look for 15 sec. at a card with a dull colored figure on it and then he was to attempt to make the best color match possible from memory. He was informed that when the sliding door was raised he would see a rotating colored disc of a color wheel and that he could instruct *E* to add more red or more yellow as the case required. Through a series of such adjustments he was to achieve the most satisfactory match he could.

For use in the color-matching series enlarged (3 ×) replicas of the six forms were made from neutral gray paper (Stoelting No. 19), and each was mounted on a separate 6 × 8-in. blue-green background (Stoelting No. 10) such that a dull orange contrast color was produced. Each card was covered with slightly opaque tracing paper to enhance the contrast effect. In addition, the shade in the experimental room was drawn and artificial lighting from a ceiling fixture was used to aid the illusion.

The color match for a given form was made from memory immediately following the 15-sec. exposure period. Each time *S* instructed *E* to adjust the rotating disc the sliding door was closed, the color wheel stopped, and an appropriate small change in proportionate sizes of red and yellow sectors was made. At no time was *S* allowed to observe the stationary wheel to see the actual sizes of the sectors. In cases where the "correct" matching position was exceeded, an adjustment in the reverse direction was allowed. All final matches were recorded in terms of the percentage of yellow used in the red-yellow mixture.

In order to control for errors of habituation and expectation, *Ss* were divided randomly into two subgroups, Group A and Group B, and matching was started from opposite ends of the red-orange-yellow continuum for the two groups. The order of presentation of forms for all *Ss* (Groups A and B), with the color from which matching started for Group A in parentheses, was: orange labeled (red); red as attribute (yellow); yellow as attribute (red); yellow labeled (yellow); orange as attribute (yellow); and red labeled (red). Within each group the two forms associated with red (one associated with color as an attribute and one labeled with the color name only) were matched starting from red once and from yellow once. Similarly the matching for the other forms was carried out with counterbalancing for the expected direction of adjustment.

Results

Table 1 shows the means and *SD*'s of the percentages of yellow used by the groups separately and combined for each of the six forms. The data for the combined groups are plotted in Fig. 1. For both color association through color as a perceived attribute and through color labeling alone, the combined groups used the highest percentage of yellow to match forms formerly associated with yellow, the next highest percentage to match forms associated with orange, and the smallest percentage to match forms associated with red. The same pattern was seen in Groups A and B with the exception of a relatively small reversal for Group B on the color-labeled forms. Except for the form

TABLE 1

Means and *SD*'s of Percentages of Yellow Used to Match Constrast Color of Forms Previously Associated with Colors in One of Two Ways

Group	Color as Attribute						Color Labeled					
	Yellow		Orange		Red		Yellow		Orange		Red	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
A	70.31	18.22	69.00	20.53	68.08	17.60	78.15	17.46	66.31	18.87	58.07	21.62
B	77.54	15.43	53.69	21.37	46.23	20.38	62.00	21.95	65.08	14.88	56.38	25.84
Both	73.92	17.29	61.35	22.44	57.15	21.97	70.08	21.39	65.69	17.02	57.23	23.82

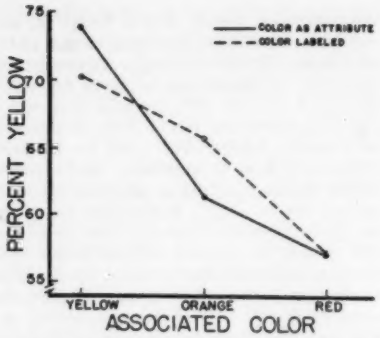

FIG. 1. The percentage of yellow in a red-yellow mixture judged to match a dull orange contrast color of forms which had previously been associated with yellow, orange, or red. Each point is an average for 26 *Ss*.

which had had red as an attribute, Group A used more yellow on the average in its matches than did Group B. No appreciable or consistent differences appeared to exist between the two types of color association in the absolute amounts of yellow employed in the color matches.

To assess the statistical significance of the results, two parallel analyses of variance were undertaken. A summary of these analyses is presented in Table 2. Considering first the results for the forms for which color was associated as an attribute, the variance attributable to colors, using the residual as the error term, is significant ($p < .01$). However, the generality of this result is conditioned by the fact that the Color × Group interaction was also significant, and the Between Colors variance was not shown to be significant over and above this interaction. With the mean square for Between *Ss* within the same group as the error term, the Between Groups variance was not significant. However, the variance Between *Ss* within the same group was significant beyond the 1% level when the residual mean square was used as the error term, thus indicating significant individual differences among *Ss*.

In the data on matching of forms which were merely labeled with color names, only one important difference in the results of the two analyses is revealed. In this case, the Color × Group interaction was insignificant. Consequently, the interaction variance was pooled with the residual variance to obtain a total residual based on 50 *df* for use as an error term. As when color was associated as an attribute of the forms, both Between Colors and Between *Ss* within the same group were shown to be significant sources of variance, although the level of confidence is somewhat less. Again, Between Groups variance indicates that the starting point of adjustments for color matches is not a significant variable.

The analyses of variance attest to the fact that the three sets of color matches for each type of color association could not have been drawn from the same population by chance alone. Figure 1 illustrates the similar general downward trends of the relationships for the two types of association procedure. For the purpose of evaluating the significance of the differences between the three points on each individual curve, the appropriate residual mean square provides the best estimate of error variance. With this error term, a difference of 5% or more in the amount of yellow used in matches is needed to indicate significance at the 5% level when color is associated as an attribute of form. A difference of 8% or more indicates significance at the 1% level. The comparable treatment of data for forms which were only labeled shows that differences of 6% or more are required to reach the 5% level of confidence and 8% or more for the 1% level. Thus, yellow-orange and yel-

low-red differences are significant at the 5% level or better for color associated as an attribute of form, while orange-red and yellow-red differences for forms associated with color by labels alone are significant. However, the present data do not allow for a more refined description of the shapes of the relationships between amount of yellow needed for matching forms and the previous color associates of the forms.

Discussion

The results in this experiment indicate that attributive color judgments do not depend solely upon wave length of stimulus light, size and brightness of stimuli, or adaptation conditions of the eye. Forms with identical physical energy characteristics are judged partially upon the basis of past color associations with the shape of the forms. In the present case color labeling alone, i.e., associations with verbal symbols for colors, was capable of exerting an influence upon future perceptions, and this type of association appeared to be as effective as when both perceived color and color name associations were combined. Apparently wide differences between associated colors are not necessarily required for the effects of previous learning to be detected in color matches.

Of particular interest is the finding that as few as 14 pairings were sufficient to "build in" color associations strong enough to affect perceptions. The uniqueness of the forms used precludes the assumption of the existence of color associates prior to experimental treatment. It is noteworthy that no *S* made any erroneous anticipations in the last 16 of the 84 exposure trials, and many *S*s had ceased making errors several trials before that. As far as the verbal learning in the pretest situation was concerned, 84 trials permitted a fair degree of overlearning. Undoubtedly this finding is dependent upon the fact that clearly distinguishable forms were employed and that only three different responses were required. Had similarity factors been different and entailed greater similarity among forms, generalization influences would have been expected to render differentiations more difficult, and, consequently, learning would have been expected to require more trials. No indication could be found in the data to signify greater difficulty in learning associations between gray forms and color names than in learning associations between colored forms and color names. Apparently previous experience had conditioned perceived color and symbolized color to be functionally quite equivalent.

The present case for the relatively minor role of motivational factors in producing perceptual effects by past color associations is admittedly weak. The pretest situation contained all the reward opportunities to be found in many human learning experiments, and undoubtedly correct anticipation of a color name was accompanied by reduction in some form of secondary motivation. However, no attempt was made to highly motivate *S*s and the situation would not be justifiably described as very ego-involving. The most that can be said on the issue of motivational influences in the present experiment is that no depend-

TABLE 2

Summary of Analyses of Variance

Source	*df*	Mean Square	*F*	*p*
Color as attribute				
Between groups	1	1940.01		
Between *S*s within same group	24	840.75	4.891*	<.01
Between colors	2	1980.17	11.518*	<.01
Colors × Group	2	1512.55	8.789*	<.01
Residual	48	171.92		
Total	77			
Color labeled				
Between groups	1	788.51		
Between *S*s within same group	24	904.50	3.865†	<.01
Between colors	2	1108.67	4.737†	<.05
Color × Group	2	468.05		
Residual	48	245.12		
Total residual	50	235.04		
Total	77			

* Residual used as error term.
† Total residual used as error term since interaction variance was not significant.

ence upon primary drive reduction was demonstrated.

While there is little doubt about the rejection of the null hypothesis concerning differences between matches for forms with different past color associations, certain restrictions upon the generality of the present results need to be recognized. In one analysis of variance there was indication of the dependence of the obtained effect upon interactions between color associates and the conditions of the measurement of color matches. Already mentioned is the possible difference in findings which might accompany variation in the particular degree of similarity among forms. Furthermore, this study tells us nothing about color matches not made from memory. As Postman (3) has pointed out, memory traces can be considered as functionally similar to small amounts of stimulus information in direct perception, and, as such, experience factors are particularly likely to reveal themselves under these conditions. Further, this investigation can shed no light upon the question of whether or not specific responses must be made to stimuli in order for past learning to affect perception of the stimuli in the future. The results of Postman and Conger (cited in 1) would indicate that response must accompany stimulation to affect recognition thresholds of stimuli presented tachistoscopically. Here verbal responses were involved, although perhaps not necessarily as a requirement for the perceptual influences to be shown. Possible differences between perceived colors and written color names in evoking verbal responses (perhaps implicit) deserve note. Finally, it is pertinent to point out that further study is needed to yield information concerning retention of associations as they affect color judgments. In the present experiment the effect of a relatively few past color associations has been shown to influence attributive color judgments made almost immediately following learning.

Summary

In order to investigate the hypothesis that attributive color judgments are functions of past color associations, 26 *S*s were shown six nonsense form-color name pairs 14 times each and then were asked to match from memory the color of each form as it appeared when shown as a gray silhouette with a dull orange contrast color. In the pretest situation three associations involved labeling alone for neutral gray forms; the other three included color as a perceived attribute of the forms. Color matches were recorded in terms of the percentage of yellow in a red-yellow mixture on a color wheel.

1. Color matches for forms associated with yellow, orange, and red required decreasing amounts of yellow in that order. Color labeling alone produced the same effects as labeling plus perceived color.

2. Analyses of variance showed the differences between matches to be statistically significant. Only two comparisons between matches for pairs of forms failed to show significance.

3. Significant individual differences between matches for different *S*s were found when color was an attribute of forms.

4. The results support the hypothesis that attributive color judgments depend upon past color associations as well as physical energy factors and adaptation conditions of the eye.

References

1. Brown, J. S., Harlow, H. F., Postman L. J., Nowlis, V., Newcomb, T. M., & Mowrer, O. H. *Current theory and research in motivation: a symposium.* Lincoln: Univ. of Nebraska Press, 1953.
2. Bruner, J. S., Postman, L., & Rodrigues, J. Expectation and the perception of color. *Amer. J. Psychol.*, 1951, **64**, 216–227.
3. Postman, L. Toward a general theory of cognition. In J. H. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads.* New York: Harper, 1951.
4. Solomon, R. L., & Postman, L. Frequency of usage as a determinant of recognition thresholds for words. *J. exp. Psychol.*, 1952, **43**, 195–201.
5. Vernon, M. D. *A further study of visual perception.* New York: Cambridge Univ. Press, 1952.

(Received June 21, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

IMPLICIT VERBAL CHAINING IN PAIRED-ASSOCIATE LEARNING[1]

WALLACE A. RUSSELL AND LOWELL H. STORMS

University of Minnesota

The role of covert symbolic processes in behavior determination has been emphasized by many psychological theorists, but the relative lack of supporting experimental evidence for this influence has been noted by several writers (**2; 8,** p. 110). Among the various factors deemed relevant to the operation of mediational or symbolic processes, the verbal habits of the individual have been prominently suggested. The work of Foley and Cofer (**3**) on verbally mediated generalization, and of Bousfield (**1**), Jenkins and Russell (**6**), and others working with recall, have established that the influence of such verbal associations can be studied effectively under experimental conditions.

While the mediational role of word associations has been demonstrated in several performance situations, there are surprisingly few studies which report positive findings with regard to the effects of mediated associations upon learning. Bugelski and Scharlock, using paired-associate learning and nonsense syllables, have provided what they term a "reasonably clear-cut demonstration of mediated association in the learning of verbal material" (**2,** p. 337). Their *Ss* showed facilitated learning of A-C associates when another term, B, presumably intervened between A and C as a consequence of previous learning of associations A-B and B-C. The term B, then, provided an implicit common term which was elicited by A and which in turn tended to elicit C. This mediation of the correct response C occurred even though *Ss* did not report deliberate use of the common term as a mnemonic device. The Bugelski and Scharlock data provide a much clearer instance of mediation than did the earlier experiments of Peters (**12**). Although the latter obtained some positive results, the majority of his test situations failed to demonstrate mediational effects.

Both Bugelski and Scharlock and Peters worked with associations learned within the context of the experiment and neither considered situations involving more than one intervening term. Nevertheless, applications of the mediation hypothesis have frequently referred to existing language habits and have almost always involved reference to a "chain" of several intervening terms which are linked on an associative basis. Uniprocess theorists who, according to Harlow (**4,** p. 452), maintain that "thinking is dependent only upon the formation and appropriate elicitation of a vast number of simple associations" have most commonly assumed that (*a*) mediational effects can occur across several intervening terms, and (*b*) these influences are at least as strong for associations between real words as between nonsense syllables. Hull's concept of pure stimulus acts (**5**), Miller's extension of the notion of verbally mediated generalization (**10,** p. 181), and Osgood's discussion of thinking (**11,** p. 638) are but a few examples in which one or the other of these assumptions has been made.

[1] This study is part of a larger series of studies of verbal behavior being conducted at the University of Minnesota. The series is being sponresod by the Office of Naval Research (Contract No. N8 onr-66216) under its policy of encouraging basic research.

However, neither Peters nor Bugelski and Scharlock provide a basis for these assumptions.

The purpose of this experiment was to test the adequacy of the above assumptions by observing the effects of mediating verbal processes on paired-associate learning when the mediating process is implemented in part by pre-existing language habits and extends over more than one implicit verbal term.

Method

Design.—The design of this study was similar to that of Bugelski and Scharlock (2). Their *Ss* learned three paired-associate lists: List 1 established A-B associations; List 2, B-C associations; and List 3 tested for mediation effects in the learning of A-C associations. In the study reported here, real words were used rather than nonsense syllables, and two implicit terms rather than one linked the pairs learned on the test trials. Here, the learning of List 1 established A-B associations. The B term was in each case a stimulus word from the Kent-Rosanoff association test (7). Recently obtained norms for responses on this test[2] then made it possible to infer certain B-C associations without establishing them experimentally. Similarly, other unpublished studies provided normative data concerning the most frequent associative responses (D) made to the C terms. Thus, once the A-B associations were learned, it was possible to infer an associative chain leading from A to B to C to D. The test for mediational effects was made by requiring *Ss* to learn a list containing A-D pairings and appropriate control pairings (A-X) of nonchained terms. It was hypothesized that the A-D pairings would be learned more easily than the A-X pairings. The manner in which associative chains might facilitate the elicitation and learning of A-D pairs is schematically illustrated as follows:

	List 1	*Associations Inferred from Norms*	*List 2 (Test List)*
Chaining Paradigm	$A_1 \cdots > B_1$	$(B_1 \rightarrow C_1 \rightarrow D_1)$	$A_1 \cdots\cdots\cdots > D_1$ $\searrow (B_1 \rightarrow C_1) \nearrow$
Control Paradigm	$A_2 \cdots > B_2$	$(B_2 \rightarrow C_2 \rightarrow D_2)$	$A_2 \cdots\cdots\cdots\cdots > X_2$ $\searrow (B_2 \rightarrow C_2 \rightarrow D_2)$

Broken arrows indicate the association to be learned in each list. Solid arrows represent associations established before a list is learned. It can be seen that associations existing before the learning of the test list provide an indirect linkage of the A and D terms which must become associated in the chaining paradigm. In the control paradigm, a similar linkage exists between A and D, but no such connection can be inferred between the A-X pairs which are to be learned.

Construction of paired-associate lists.—Tables 1 and 2 contain the associative frequencies from the norms[2] for the three most frequent responses to the words relevant to this experiment. Table 1 provides this information for the ten Kent-Rosanoff stimulus words employed and Table 2 provides analogous data for the most frequent responses to those ten words.

Table 3 contains the nonsense syllables, the particular associative chains, and the control words used throughout the experiment. List 1 (A-B pairings) was made up of the nonsense syllables in Column A and the corresponding words in Column B. Columns B, C, and D list the verbal associative chains ($B \rightarrow C \rightarrow D$) derived from the normative tables. Two test lists were formed. One combined the first five nonsense syllables of Column A with the five corresponding words of Column D and the last five nonsense syllables with the corresponding words of Column X. The other combined the first five nonsense syllables in Column A with the five corresponding words of Column X and the last five nonsense syllables with the five corresponding words of Column D. Thus the two lists were counterbalanced and each contained five A-D and five A-X pairs. The response words of Form 1 of the test list are followed by a "1" in the table; the remaining words made up Form 2.

The ten nonsense syllables of Column A were selected from Melton (9), and all had Glaze association values of 0%. The ten verbal chains listed in Columns B, C, and D of the table were selected so that as far as the complete norms would indicate, no word in any chain ap-

[2] Revised norms for 100 words from the Kent-Rosanoff word-association test were obtained from 1008 students in beginning psychology classes at the University of Minnesota. This work was carried out as part of a larger project on verbal behavior. Information concerning these norms may be obtained from the authors.

[3] See footnote 2.

TABLE 1

ASSOCIATIVE FREQUENCIES OF THE THREE MOST FREQUENT RESPONSES TO TEN KENT-ROSANOFF STIMULUS WORDS
(Based on R's from 1008 *Ss*)

Stimulus	Primary R	*f*	Secondary R	*f*	Tertiary R	*f*
STEM	FLOWER	402	PLANT	224	LEAF	125
MEMORY	MIND	119	REMEMBER	99	FORGET	80
SOLDIER	ARMY	187	SAILOR	182	MAN	101
TROUBLE	BAD	89	SHOOTER	49	WORRY	45
WISH	WANT	124	DREAM	118	DESIRE	112
JUSTICE	PEACE	250	LAW	182	COURTS	163
THIEF	STEAL	286	ROBBER	138	CROOK	69
OCEAN	WATER	314	SEA	233	BLUE	111
COMMAND	ORDER	196	ARMY	102	OBEY	78
FRUIT	APPLE	378	VEGETABLE	114	ORANGE	94

peared among the ten most frequent responses to any word in any other chain. In addition, no final word (D) in a chain appeared as a response to the first word (B) more than seven times in the 1008 responses to that word in the norms. It may be noted that while Column C is essential to the construction of the ten associative chains, no words in that column appeared in any of the experimental lists. The control words in Column X were chosen from the Kent-Rosanoff list such that none appeared among the ten most frequent responses to any of the chained words and no chained word appeared among the ten most frequent responses to a control word. Finally, the control words were matched with the final words of each chain on the basis of Thorndike-Lorge (13) frequency as a partial equalization of difficulty between control and experimental words.

Procedure.—Twenty-seven sophomore women from a beginning psychology class served as *Ss*.

Following general instructions on paired-associate learning, each *S* learned List 1 on a standard Hull-type memory drum. Each stimulus word was exposed for 2 sec. before the response word appeared beside it for another 2-sec. period. The next stimulus word followed immediately, except that 4 sec. elapsed after each complete trial through the list. The *S* was required to learn the ten pairs in List 1 to a criterion of three consecutive trials in which all response words were correctly anticipated. To control for serial position effects, the list was presented in three successive random orders of

TABLE 2

ASSOCIATIVE FREQUENCIES OF THE THREE MOST FREQUENT RESPONSES TO TEN PRIMARY RESPONSES TO KENT-ROSANOFF STIMULUS WORDS
(Based on R's from 100 *Ss*)

Word	Primary R	*f*	Secondary R	*f*	Tertiary R	*f*
FLOWER	SMELL	15	ROSE	12	PRETTY	12
MIND	BRAIN	15	MATTER	14	THINK / SOUL	12
ARMY	NAVY	39	SERVICE	7	SOLDIER / MAN	4
BAD	GOOD	71	EVIL	4	MEAN	3
WANT	NEED	27	DESIRE	19	HAVE	8
PEACE	WAR	42	DOVE	10	QUIET	8
STEAL	TAKE	14	THIEF	10	ROB	8
WATER	DRINK	19	THIRSTY	14	WET	11
ORDER	DISORDER	14	COMMAND	6	STOP	5
APPLE	TREE	16	RED	14	ORANGE	13

TABLE 3

NONSENSE SYLLABLES, ASSOCIATIVE CHAINS, AND CONTROL WORDS USED IN FORMING THE PAIRED-ASSOCIATE LISTS

A	B	C	D	X
Nonsense Syllable	First Chained Word	Second Chained Word	Final Chained Word	Control Word
CEF	STEM	FLOWER	SMELL (1)	JOY
DAX	MEMORY	MIND	MATTER (1)	AFRAID
YOV	SOLDIER	ARMY	NAVY (1)	CHEESE
VUX	TROUBLE	BAD	GOOD (1)	MUSIC
WUB	WISH	WANT	NEED (1)	TABLE
GEX	JUSTICE	PEACE	WAR	HOUSE (1)
JID	THIEF	STEAL	TAKE	SLEEP (1)
ZIL	OCEAN	WATER	DRINK	DOCTOR (1)
LAJ	COMMAND	ORDER	DISORDER	CABBAGE (1)
MYV	FRUIT	APPLE	RED	HAND (1)

Note.—The words and syllables were presented in capital letters exactly as above. The response words of Form 1 of the test list are followed by a "1."

pairs before the first order was repeated. Three *Ss* failed to reach the criterion on List 1 within 40 min. and were not used further in the experiment.

After a pause of 4 min., each remaining *S* was presented with either Form 1 or Form 2 of the test list. The *Ss* were instructed that the procedure was exactly the same as for List 1 and were urged to do their best on the test list. Since one *S* failed to reach the criterion of one trial in which all the response words were correctly anticipated, there remained 23 *Ss* for the final analysis of results. Twelve of these learned Form 1 of the test list and 11 learned Form 2.

For each *S*, all correct anticipations and errors were tabulated for both List 1 and the test list. Measures used in the final analysis were: (*a*) the number of trials required to reach the criterion on List 1; (*b*) the number of mediated (D) and unmediated (X) responses occurring in the first five different correct anticipations by each *S*; (*c*) the total number of correct responses made for the mediated (D) and unmediated (X) words during the test trials.

Control experiment.—Twelve additional female *Ss* from the same population were used in a subsequently performed control experiment designed to allow a comparison of the learning of A-D and A-X pairs in a situation where associative chaining could not differentially contribute to the learning of the pairs. Instead of learning List 1, these *Ss* first learned either Form 1 or Form 2 of the test list. The second list was the other form of the test list. Thus, in the control experiment, associative facilitation due to A → B → C → D linkages was not possible because the A-B associations of List 1 were not learned by any *S*. The analysis of results for the control *Ss* was based on measures similar to those of the main experiment.

RESULTS

Since two forms of the test list and two groups of *Ss* were involved in the design, the equivalence of the two lists and the two groups must be established before the combined results can be dealt with. An analysis of the mean number of trials required to reach the criterion of learning on List 1, which all *Ss* learned, indicated that the 12 *Ss* who subsequently learned Form 1 of the test list did not differ significantly ($t = .52$) from the 11 *Ss* who subsequently learned Form 2 of the test list. The mean for the former group was 25.00 ($SD = 9.17$); for the latter it was 22.91 ($SD = 9.30$). With respect to performance on the two forms of the test list, the mean number of correct anticipations per *S* during learning did not differ significantly between Forms 1 and 2. The means for Forms 1 and 2 were, respectively, 67.08 ($SD = 30.75$) and 58.09 ($SD = 17.04$). The Behrens-Fisher d of .87 did not allow rejection of the null hypothesis concerning form differences.

As a consequence of these comparisons, the results from Forms 1 and 2 of the test list were combined. In order to determine whether the response terms for A-D pairs, for which associative chaining was possible, were more easily elicited during the early trials, an analysis was made of the first five different correct responses made by each *S*. Of the 115 responses in this analysis, 67 were members of "chained" A-D pairs and 48 were from "unchained" A-X pairs. The normal curve approximation to the binomial indicates that a result this large and in this direction would occur by chance less than 4 times in 100 if the probabilities of successes for A-D and

A-X pairs were equal. Of the 23 *Ss* in this part of the experiment, 19 showed more chained than unchained responses in their first five different correct responses. The same one-tailed binomial test indicated that this result would occur by chance only 1 time in 1000 if the probabilities of successes for all pairs were equal. The conclusion that chained words were more easily elicited during the early trials of learning seemed warranted.

The major purpose of the experiment, however, was to compare the ease with which chained and unchained pairs were learned. The design allowed this comparison to be made with each *S* acting as her own control. Since each *S* learned an equal number of chained (A-D) and unchained (A-X) pairs, the total number of correct anticipations by each *S* for the unchained pairs was subtracted from the corresponding total for chained pairs. If there is facilitation of the learning of chained pairs (i.e., *S* has a larger number of correct anticipations on the A-D than on the A-X pairs) this difference will be positive. Over all *Ss* the mean difference between chained and unchained pairs was 3.74 ($SD = 5.32$). A t of 3.30 ($.01 > p > .001$) leads to rejection of the null hypothesis, and the conclusion that there was facilitation of learning of A-D pairs as contrasted with the learning of A-X pairs.

A secondary analysis of performance on Form 1 and Form 2 separately revealed that the direction of the difference between chained and unchained pairs was positive for both forms (Form 1, $M = +4.92$, $SD = 5.85$; Form 2, $M = +2.45$, $SD = 4.31$). For Form 1, the t of 2.79 was significant at the .02 level of confidence. For Form 2, the t of 1.80 was between the .15 and .10 levels.

TABLE 4

SUMMARY OF MAJOR RESULTS

Experiment	Initial Successful Anticipations (First 5 Different Correct R's for Each *S*)			Total Correct Responses: Mean of Chained Minus Unchained	
	Chained	Unchained	Significance of Difference	(3)	Significance
Main	67	48	.04	+3.74	.01
Control	30	30		−0.50	

It was recognized that, if for reasons other than associative chaining, the A-D pairs were as a group intrinsically easier to learn than the A-X pairs, the results obtained here could be accounted for on the basis of that factor alone. The control study was run to provide information about the relative difficulty of A-D and A-X pairs in a situation where chaining of A-D pairs was not possible. The 12 *Ss* in the control experiment learned both Forms 1 and 2 of the test list. The performance of these *Ss* on whichever form was learned last provided the basis for the analysis of the control experiment. First, the mean number of correct anticipations of the response word was determined for the ten A-D and the ten A-X pairs. The means were 7.06 and 7.35 respectively, with *SD*'s of 1.29 and 1.39. This difference did not approach significance and the direction of difference is unfavorable to the hypothesis that the A-D pairs were easier to learn than the A-X pairs. Finally, the two major analyses of the main experiment were repeated here. As Table 4 indicates, neither of the differences tested was significant, and in each case the direction of difference did not favor the A-D pairs. In the absence of the possibility of associative chaining, then, there was no evidence of easier

learning of the A-D pairs used in this experiment.

Discussion

Statistically, these results provide stronger evidence for mediational effects in learning than do the results of Bugelski and Scharlock (2). This is true in spite of the fact that the present experiment involves one more step in the chain of associations mediating the facilitated learning. Instead of an A-B-C sequence contributing to the learning of A-C, an A-B-C-D chain contributed to the learning of A-D. This demonstration of mediational influences extending over more than one intervening term, and involving language habits established prior to the experiment, offers some confirmation for theoretical explanations of thinking, problem solving, etc., which have postulated the operation of such complex implicit associative sequences.

Of course, the highly significant results obtained here, in the face of less convincing evidence obtained in schematically simpler situations (2), raise the problem of accounting for this stronger effect. Two possibilities occur to the writers. First, it is probable that this experiment allowed a more efficient analysis by removing variability due to individual differences in learning ability. Although Bugelski and Scharlock endeavored to have each *S* act as his own control, their technique of analysis admittedly left some individual difference factors operating. Our procedure of using within-individual differences removed this variable and may have allowed mediational effects to be revealed more sensitively. Furthermore, it is at least conceivable that the pre-existing verbal habits of this experiment were stronger than the associations learned during the Bugelski-Scharlock experiment. It is probable that such strong associations, if such they were, brought about mediational effects more readily than weaker associations would have done.

The mere demonstration of mediational influences in learning, however, does not explain how the effect is achieved. The most plausible explanation would hold that the presence of an associative chain between the stimulus term and the response term in paired-associate learning increases the probability that the response term will be elicited in the learning situation. Any such elicitation would presumably have two effects. First, it would increase the total number of correct responses made during learning. This would be a performance change influencing the criterion measures used in this study. Second, there would be an influence on learning. The performance change, of course, does not necessarily reflect a change in the underlying learning process itself. Nevertheless, such a learning change is implied, since any factor which increases the frequency of occurrence of a correct response would increase the number of reinforced trials and thus indirectly influence the amount of learning.

Less obvious is the possibility that the differences between the mediated and unmediated pairs are due to interference effects in the learning of the control (A-X) pairs. Although interference due to the tendency for the A terms to elicit B was controlled by the design, possible differential interference effects may be seen when the entire A-B-C-D sequence is considered. If the probability of the elicitation of D is enhanced by the presence of A, as is stated above, then this tendency would compete with the elicitation of the correct response X in the unmediated pairs and possibly delay learning. It is conceivable that the associative chains used here produced both a facilitative effect upon mediated pairs and an interference effect upon unmediated pairs. The possibility that these two effects of associative chains do operate is amenable to experimental test, although the design of this experiment and that of Bugelski and Scharlock (2) do not allow an analysis which would separate them.

Whatever the explanation of the mediational effect may be, there can be little doubt that it is the phenomenon underlying the superior performance of *S*s on

the A-D pairs. The controls inherent in the main experiment plus the additional information from the control experiment leave little room for alternative hypotheses. Such factors as serial position, idiosyncrasies of words and *S*s, etc. operated equally for the mediated and nonmediated pairs and could not account for the differences obtained.

The fact that questions following the experiment yielded no evidence that *S*s could verbalize the mediating terms only emphasizes the Bugelski-Scharlock conclusion that mediated association may be "unconscious."

Summary

This experiment was designed to study the effects of mediating verbal processes on paired-associate learning when the mediating process is implemented in part by pre-existing language habits and extends over more than one implicit verbal term.

First, ten chains of word associations, B-C-D, were constructed from normative data on association frequencies. Twenty-three female college *S*s then learned a list of A-B pairs where the A terms were nonsense syllables and the B terms were the initial members of the chains described above. The test situation required that *S*s learn another list consisting of A-D and A-X pairs. The D terms were the final members of the associative word chains, and the X terms were not associated with any of the chains. A control experiment revealed that the A-D and A-X pairs did not differ in difficulty in the absence of chaining possibilities.

It was found that the A-D pairs were learned significantly faster, and elicited earlier in learning, than the A-X pairs. It was concluded that implicit verbal chains of more than one link mediated these effects. Reasons for these results being even more clear-cut than those of schematically simpler previous experiments were discussed.

References

1. Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates. *J. gen. Psychol.*, 1953, **49**, 229–240.
2. Bugelski, B. R., & Scharlock, D. P. An experimental demonstration of unconscious mediated association. *J. exp. Psychol.*, 1952, **44**, 334–338.
3. Foley, J. P., Jr., & Cofer, C. N. Mediated generalization and the interpretation of verbal behavior: II. Experimental study of certain homophone and synonym gradients. *J. exp. Psychol.*, 1943, **32**, 168–175.
4. Harlow, H. F. Thinking. In H. Helson (Ed.), *Theoretical foundations of psychology*. New York: D. Van Nostrand, 1951. Pp. 452–500.
5. Hull, C. L. Goal attraction and directing ideas conceived as habit phenomena. *Psychol. Rev.*, 1931, **38**, 487–506.
6. Jenkins, J. J., & Russell, W. A. Associative clustering during recall. *J. abnorm. soc. Psychol.*, 1952, **47**, 818–821.
7. Kent, G. H., & Rosanoff, A. J. A study of association in insanity. *Amer. J. Insanity*, 1910, **67**, 37–96, 317–390.
8. McGeoch, J. A., & Irion, A. L. *The psychology of human learning*. New York: Longmans, Green, 1952.
9. Melton, A. W. Materials for use in experimental studies of the learning and retention of verbal habits. Unpublished manuscript, Univ. of Missouri, 1940.
10. Miller, G. A. *Language and communication*. New York: McGraw-Hill, 1951.
11. Osgood, C. E. *Method and theory in experimental psychology*. New York: Oxford Univ. Press, 1953.
12. Peters, H. N. Mediate association. *J. exp. Psychol.*, 1935, **18**, 20–48.
13. Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words*. New York: Teachers Coll., Columbia Univ., 1944.

(Received June 26, 1954)

Journal of Experimental Psychology
Vol. 49, No. 4, 1955

THE RELATIONSHIP BETWEEN OCCIPITAL ALPHA ACTIVITY AND LATERALITY[1]

A. D. GLANVILLE AND J. J. ANTONITIS

University of Maine

The purpose of this study was to determine if differences in measures of alpha activity between the right (R) and left (L) occipital areas in normal adult *Ss* are correlated with laterality or sidedness as measured by the Jastak Ambigraph Laterality Test. Although marked differences in electrical activity between homologous regions on the two sides of the head are considered indicative of pathological conditions, small differences are not unusual in normal *Ss* (**2, 3, 11**). Some writers in this area believe that these "normal" differences are related to the handedness or laterality of the individual (**2, 9**).

Raney (**9**) in a careful quantitative study of the relationship between alpha activity and lateral dominance in a group of 17 pairs of identical twins between 7 and 16 yr. of age found " . . . a tendency for the non-dominant side of the head according to behavioral indices to have a greater amount and amplitude of alpha activity" (p. 38). He found that such bilateral differences were more apparent in individuals with very pronounced lateral dominance, but the differences between the averages for his R-dominant and L-dominant *Ss* were not significant statistically "due to the small number of cases."

In 1940 Lindsley (**7**) reported a study of the relationship between laterality and phase relationships of alpha waves and the amount of unilateral blocking of the alpha on the two sides of the head in a group of 65 normal children and 2 adult stutterers. Although his results indicated that both of the alpha factors appeared to be related to laterality as well as to stuttering, he concluded that neither of these factors selectively differentiates the hemispheres in terms of dominance.

When Rheinberger, Karlin, and Berman (**10**) examined the relationship between laterality and bilateral asymmetry of electrical brain activity in ten stuttering and ten nonstuttering children, they also found no consistent relationship. They pointed out, however, that there was a tendency for a greater degree of bilateral asymmetry to occur in individuals who showed strong lateral dominance on the laterality tests used.

A more recent statement of the relationship between bilateral differences in EEG activity and laterality is that of Cohn (**2**). Apparently largely on the basis of impression gained from the clinical evaluation of a number of normal EEG records he states: "In approximately half of all subjects recorded from, there is a definitely discernible amplitude asymmetry over each cerebral hemisphere. In the great majority of these cases with amplitude asymmetry the output from the right cerebral hemisphere is more rhythmic and of higher voltage than that from the left side. This in a crude way corresponds to dominant handedness or eyedness, but there is no strict relationship between handedness and electrical cerebral dominance" (**2**, p. 9).

This brief survey of studies and statements which bear directly on the problem indicates that further quantitative study of the relationship be-

[1] The substance of this paper was presented at the 1954 meetings of the Eastern Psychological Association in New York City. The research was made possible by a grant to the authors from the Wilson Coe Research Fund. Appreciation is expressed to Joseph J. Majors and Stanley L. Payson, Jr., graduate students in psychology at the University of Maine, for invaluable assistance in the conduct of this study.

tween bilateral differences in alpha activity and laterality is needed. The problem may best be dealt with by making careful quantitative measurements both of alpha activity and of laterality on a relatively large group of *Ss*. The present writers, using data obtained in a normative EEG survey of adult *Ss*, were provided with an opportunity to examine carefully the possible relationship between bilateral differences in occipital alpha activity and sidedness.

Since a possible weakness of previous studies may have been the method of determining laterality of the *Ss*, the Jastak Ambigraph Laterality Test (4) was chosen for use in the present study. Its author indicates that the test eliminates the effect of training to a greater degree than many other tests of laterality. The test has the additional advantages of not being restricted to some single aspect of laterality and of providing percentile norms established on 820 individuals.

Method

Subjects.—Fifty normal adult male *Ss* were used. The group included students and teaching staff members of the University of Maine. The mean age of the group was 23.8 yr.; the *SD* was 5.5 yr.; and the range was 19 to 47 yr. Eighty-six per cent of the *Ss* in the group were under 28 yr. All *Ss* were volunteers; the only criterion used for selection was willingness to serve as *Ss*.

Normality of the *Ss* with respect to factors which might conceivably influence EEG's was established by questioning *Ss* at the time of examination. None of the *Ss* in the group reported having suffered severe head injury, or psychomotor attacks which might be indicative of epilepsy, or any form of severe psychosomatic disorder.

Apparatus and procedure.—A Model III-DS, two-channel Grass electroencephalograph was used in making simultaneous monopolar recordings of electrical activity in the R and L occipital areas of each *S*. The *S* was seated in a lounge chair in an electrically shielded room. Silver coil electrodes (5) were fixed to the scalp at the level of the inion and positioned laterally on each side of the inion between 1½ and 2 in. In all instances the electrodes were placed within R and L areas of maximal alpha activity as specified by Adrian and Yamagiwa (1). Silver ear clips were used as reference and ground electrodes on R and L ear lobes, respectively. Electrode jelly was used in the application of the electrodes. The resistance between the occipital electrodes and the reference electrode as measured with an ohmmeter was 10,000 ohms or less in every instance both prior to and following the EEG examination. After the electrodes were attached, the room was darkened and *S* was instructed to close his eyes and to relax as fully as possible.

EEG activity was recorded for 4 min. with a pause of 30 to 60 sec. between the first and the second 2-min. periods. During the pause the EEG operator entered *S*'s room to make minor electrode adjustments if necessary, to reassure *S*, and to repeat the instructions to keep the eyes closed, to remain immobile, and to relax as fully as possible. The tape speed was 30 mm. per second. Both EEG channels were calibrated so that 1 mm. of pen deflection represented an amplitude of 10 microvolts. Immediately after the EEG recording, *S* was given the Jastak Ambigraph Laterality Test.

Analysis of records.—Since the method adopted for measurement of aspects of alpha activity was painstaking and extremely time consuming, it was decided to measure the last 40 sec. of record free of artifacts at the end of the second 2-min. period. This portion of the record was chosen on the assumption that it would provide the most reliable indication of electrical activity since *S* would be expected to be somewhat more adapted to the situation at the end than at the beginning of the examination. The alpha activity in the 40-sec. period selected was then marked off for measurement. The criteria adopted for determining what waves were to be considered as alpha were essentially those of Shagass (12), namely, that there should be at least three consecutive waves with a frequency of 7 to 14 per second, and that each wave should have an amplitude of not less than 15 microvolts. A special clear plastic gauge with hairline separations etched to indicate these standards was employed to insure uniformity of definition of alpha activity.[2]

After sections of the record containing alpha were marked off, these sections were measured with a millimeter scale to determine the amount of alpha and the alpha index (percentage of alpha) for each lobe. All amplitude measure-

[2] The gauge used was constructed by Professor Jonathan Biscoe of the Department of Physics of the University of Maine.

ments were made with an "alpha cumulator" similar to one devised by Shagass and Johnson (13). The principles specified by Kreezer (6) for amplitude measurements were strictly adhered to.

Statistical procedures.—Since the study was concerned with the question of whether or not differences in alpha activity in the R and L occipital lobes were correlated with sidedness, it was decided to employ ratios of alpha measures in the two lobes as indices of differences in degree and direction of alpha activity and to correlate these ratios with raw scores on the Jastak Ambigraph Laterality Test.

In determining the alpha indices and mean alpha amplitudes for R and L lobes of all *Ss*, a problem arose with regard to the records of a number of *Ss* who had zero or extremely small amounts of alpha in one or both lobes. In some instances it was felt that the very small number of waves present was insufficient for an accurate appraisal of mean alpha amplitude; in other instances, the mere fact of zero values made it impossible to obtain meaningful ratios between R- and L-lobe activity. For these reasons it was decided to exclude the records of 10 *Ss* in the original group of 60. The 50 *Ss* included in the study had at least 5% alpha in one lobe and some amount greater than zero in the other.

To assess the reliability of the measurements of alpha index and mean alpha amplitude, odd-even reliability coefficients were computed using measurements obtained in alternate 10-sec. intervals of the R-lobe records of the 50 *Ss*. For alpha index, the reliability coefficient was .97; for alpha amplitude, .94. Both *r*'s have been corrected by means of the Spearman-Brown prophecy formula.[3] For the Ambigraph Laterality Test, Jastak (4) has reported that values of reliability coefficients ranged from .67 to .84 for items in the test battery in a group of 820 *Ss*. No reliability coefficient for the test as a whole is reported.

After the reliability of the measures of alpha indices and mean alpha amplitudes had been established, R/L ratios of these measures were computed for each *S*. Such R/L ratios were obtained by dividing R-lobe measures by L-lobe measures. Ratios with values less than 1.00 represented higher degrees of alpha activity in L lobes; ratios with values greater than 1.00 represented higher degrees of alpha activity in R lobes. The reliability of the R/L ratio measures of bilateral asymmetry was determined by correlating measures obtained in alternate 10-sec. intervals of the 40-sec. sections of tape measured. The reliability coefficients for R/L ratios of alpha indices and mean alpha amplitudes as corrected by means of the Spearman-Brown formula were .85 and .79, respectively.

To assess the degree of independence of these measures of bilateral asymmetry of alpha activity, the correlation between the R/L ratios of alpha indices and the R/L ratios of mean alpha amplitudes was computed. The value of *r* in this instance was .39; σ_r was .12. Even though the degree of positive correlation is significant, the relatively low value of *r* indicates that the measures are relatively independent.

Results and Discussion

The values of the correlation coefficients obtained between each of the two measures of bilateral alpha asymmetry and the laterality test raw scores, shown in Table 1, indicate no significant relationship between either of the EEG measures and sidedness. The small positive correlation between alpha index R/L ratios and laterality test scores indicates, if anything, a very slight contralateral association of greater occipital alpha activity and dominant sidedness in the present group of *Ss*. So far as occipital alpha activity is concerned, the obtained correlations do not confirm the expectation based on the previous observations of Cohn (2) and Raney (9) of greater alpha activity in the nondominant hemisphere as indicated by sidedness.

TABLE 1

Correlations Between R/L Ratios and Laterality Test Raw Scores

Measures Correlated	*r*	σ_r
Alpha index R/L ratios and laterality scores	.14	.14
Alpha amplitude R/L ratios and laterality scores	.002	.14

[3] Measurements of mean alpha frequency were also made but not used in the present study. The odd-even reliability coefficient for the alpha frequency measure corrected by the Spearman-Brown formula was .88 for the right lobe.

The results of the present study were treated in another way to obtain a direct comparison with the findings of Raney (9) concerning the direction of association of cerebral dominance as indicated by alpha activity and dominance in terms of sidedness. Alpha indices and measures of alpha amplitude reported by Raney for occipital and central regions of individual *Ss* were converted to R/L ratios, and means of such ratios were computed for Raney's R- and L-sided groups. It should be noted that Raney classified his *Ss* as dominantly R or L sided chiefly on the basis of performance on the phi test. The *Ss* of the present study were divided into R- and L-sided groups on the basis of percentile ratings of scores on the Ambigraph Laterality Test. The *Ss* scoring at or below the 50th percentile were classified as R sided, and *Ss* scoring above the 50th percentile as L sided. Means of R/L ratios for the R- and L-sided groups of the present study were then computed. Table 2 presents the means of R/L ratios thus obtained for R- and L-sided groups of the present study and Raney's study. Nonparametric tests of significance (15) were applied to distributions of R/L ratios for R- and L-sided groups to determine if the associations of side of greater alpha activity and sidedness represented in the means listed in Table 2 were statistically significant. These values are shown in Table 2.

The distributions of R/L ratios of both measures of occipital alpha activity are not significantly different for the R- and L-sided groups of both the present study and Raney's study. In both studies, the same reversal in direction of association of side of greater alpha activity and dominant sidedness is indicated by the means of the occipital R/L ratios of alpha amplitudes and alpha indices. The conclusion is that the results reveal neither a contralateral nor an ipsilateral association of greater occipital alpha activity with dominant sidedness. The present treatment of Raney's data for the central region, however, does confirm his statement based on other statistical treatments of the same data that " . . . there is a tendency for the non-dominant side of the head according to behavioral indices to have a greater amount and amplitude of alpha activity . . . " (9, p. 38). Table 2 indicates that the ipsilateral association of side of greater alpha activity measured in the central region and dominant sidedness in terms of the phi test is more clearly reflected by R/L ratios of alpha indices than R/L ratios of alpha amplitudes.

TABLE 2

COMPARISON OF MEAN R/L RATIOS OF ALPHA ACTIVITY FOR RIGHT- AND LEFT-SIDED *Ss* IN THE PRESENT STUDY AND IN RANEY'S STUDY

Measure	Present Study		Raney's Study		
	N	Occipital	*N*	Occipital	Central
Alpha amplitude R/L					
Right-sided *Ss*	22	1.001	17	1.02	1.21
Left-sided *Ss*	28	.999	17	.97	1.00
Alpha index R/L					
Right-sided *Ss*	22	1.087	17	.99	1.21*
Left-sided *Ss*	28	1.168	17	1.01	.94*

* $p < .01$ for difference between R-sided and L-sided *Ss*.

To check on Raney's observation (9) that bilateral differences in direction of greater alpha activity were more apparent in *Ss* with pronounced lateral dominance, the relationship of side of greater alpha activity to dominant sidedness was examined in *Ss* classifiable as *strong R* and *strong L* in the present study. The *Ss* scoring at or below the 24th percentile on the Ambigraph Laterality Test were classified as strong R; *Ss* scoring at or above the 76th percentile were classified as strong L. Table 3 presents the values of the means of the two types of R/L ratios for the strong R and strong L

TABLE 3

MEAN R/L RATIOS OF ALPHA ACTIVITY FOR STRONG RIGHT AND STRONG LEFT *Ss*

R/L Measure	Ambigraph Test Percentiles	
	0–24 Strong Right (N = 12)	76–100 Strong Left (N = 18)
Alpha index R/L	1.00	1.20
Alpha amplitude R/L	1.03	1.01

groups. These means suggest that the groups do not differ. The direction of the greater difference between the means of R/L ratios of alpha indices suggests that the strong L *Ss* show a tendency for more alpha activity to occur in R or dominant hemispheres as inferred on the basis of sidedness. The strong R *Ss*, with a mean R/L ratio of 1.00, do not show an opposite tendency. Nonparametric *T* tests of significance (**14**), however, reveal that the distributions of R/L ratios of both alpha indices and amplitudes are not significantly different in the strong R and strong L groups, which indicates roughly that the means do not differ significantly (**8**). The tendency noted by Raney (**9**) for a more clear-cut association of greater alpha activity in the nondominant hemisphere with dominant sidedness is not apparent in our *Ss*. The present finding, however, does not necessarily conflict with Raney's observation since Raney pointed out that the association being considered was not great enough to be statistically significant either in terms of occipital or central measures in two groups of six *Ss* each (6 twin pairs) highly selected for lateral dominance.

Thus, the data of both the present study and Raney's study indicate that such measurable differences in occipital alpha activity as do occur in simultaneous bilateral records of normal *Ss* under "relaxed" conditions are not significantly related to sidedness. Given Raney's findings (**9**) for the central region and Cohn's impressions (**2**) formed apparently on the basis of examination of EEG records taken simultaneously from several regions, bilateral differences in alpha activity recorded from central regions of relaxed *Ss* may, however, bear some relationship to sidedness. The fact that Rheinberger, Karlin, and Berman (**10**) reported that their records from central regions showed variable association of bilateral differences in alpha and sidedness indicates that further experimentation is necessary to establish conclusively that differences in alpha activity from homologous central regions are related to sidedness.

SUMMARY

Relationships between indices of differences in alpha activity recorded from right and left occipital lobes and sidedness as measured by the Jastak Ambigraph Laterality Test were examined in a group of 50 normal adult male *Ss*. No significant relationships were found either in the total group or in selected strong right and strong left subgroups. The conclusion is drawn that such measurable differences as do occur in simultaneous bilateral records of occipital alpha activity of normal *Ss* under relaxed conditions are not significantly related to sidedness.

REFERENCES

1. ADRIAN, E. D., & YAMAGIWA, K. The origin of the Berger rhythm. *Brain*, 1935, **58**, 323–351.
2. COHN, R. *Clinical electroencephalography*. New York: McGraw-Hill, 1949.
3. JASPER, H. H. Electrical signs of cortical activity. *Psychol. Bull.*, 1937, **34**, 411–481.
4. JASTAK, J. The ambigraph laterality test. *J. appl. Psychol.*, 1939, **23**, 473–487.
5. KREEZER, G. The electro-encephalogram and its use in psychology. *Amer. J. Psychol.*, 1938, **51**, 737–759.
6. KREEZER, G. Intelligence level and occipital alpha rhythm in the mongolian type of mental deficiency. *Amer. J. Psychol.*, 1939, **52**, 503–532.
7. LINDSLEY, D. B. Bilateral differences in brain potentials from the two cerebral hemispheres in relation to laterality and

stuttering. *J. exp. Psychol.*, 1940, **26**, 211–225.
8. Moses, L. E. Non-parametric statistics for psychological research. *Psychol. Bull.*, 1952, **49**, 122–143.
9. Raney, E. T. Brain potentials and lateral dominance in identical twins. *J. exp. Psychol.*, 1939, **24**, 21–39.
10. Rheinberger, M. B., Karlin, I. W., & Berman, A. B. Electroencephalographic and laterality studies of stuttering and nonstuttering children. *Nerv. Child*, 1943, **2**, 117–133.
11. Schwab, R. S. *Electroencephalography in clinical practice.* Philadelphia: Saunders, 1951.
12. Shagass, C. Conditioning the human occipital rhythm to a voluntary stimulus. A quantitative study. *J. exp. Psychol.*, 1942, **31**, 201–209.
13. Shagass, C., & Johnson, E. P. The course of acquisition of a conditioned response of the occipital alpha rhythm. *J. exp. Psychol.*, 1943, **33**, 201–209.
14. White, C. The use of ranks in a test of significance for comparing two treatments. *Biometrics*, 1952, **8**, 33–41.
15. Wilcoxon, F. *Some rapid approximate statistical procedures.* Stamford, Conn.: American Cyanamid Co., 1949.

(Received June 23, 1954)

ERRATUM

In the article "The Relationship Under Stress Between Changes in Skin Temperature, Electrical Skin Resistance, and Pulse Rate," by Lawrence M. Baker and William M. Taylor, in the November 1954 issue of this journal (*J. exp. Psychol.*, **48**, 361–365), there were two errors on page 364 in the section on Results. In the paragraph on relationships between physiological indices, the relationship for the rest period just preceding spark stimulation should be .487 (*not* .478), and the relationship under tension stimulation between skin temperature and pulse should be .348 (*not* .381). Both correlation ratios are given correctly in Table 2.